REVIEW

140

of the

TRICAL COMMUNICATION LABORATORY

PON TELEGRAPH AND TELEPHONE PUBLIC CORPORATION

Former Reports of the Electrical Communication Laboratory
Nippon Telegraph and Telephone Public Corporation
from Vol.1, No.1, Sept. 1953 to Vol. 7, No. 12, Dec. 1959

Published Bimonthly by
E ELECTRICAL COMMUNICATION LABORATORY
IPPON TELEGRAPH AND TELEPHONE PUBLIC CORPORATION
1551, Kitizyôzi, Musasino-si, Tôkyô, Japan

VOLUME 9
NUMBERS 1-2

January-February
1961

Review of the Electrical Communication Laboratory

EDITORIAL COMMITTEE

Susumu OKAMURA *Chairman*

Tadasu FUKAMI
Giichi ITO
Hideo KAWASAKI
Yukio NAKAMURA
Tatsuya SHIMBORI
Takuzo SHINDO
Toyotaro SHIRAMATSU
Takakichi UZAWA
Ginsaku YASAKI

Koji MAEDA
Nobukazu NIIZEKI
Osamu TUKIZI

EDITORIAL STAFF

Susumu HIRANO *Editor*

Tutomu SAITO *Assistant Editor*

Masao HIROSE 〃

The Review of the Electrical Communication Laboratory is published six times a year (bimonthly) by the Electrical Communication Laboratory, Nippon Telegraph and Telephone Public Corporation, 1551, Kitizyôzi, Musasino-si, Tôkyô, Japan. Telephone: Tôkyô (391) 2261 and 2271.

All rights of republication, including translation into foreign languages, are reserved by the Electrical Communication Laboratory, NTT.

Subscriptions are accepted at ¥1,200 per year. Single copy ¥300 each. Foreign postage is ¥300 per year or ¥50 per copy. Remittance should be made in check payable to the Electrical Communication Laboratory and sent to the Director.

Printed by Aoyama Printing Co., Tokyo, Japan

REVIEW
of the
ELECTRICAL COMMUNICATION LABORATORY

NIPPON TELEGRAPH AND TELEPHONE PUBLIC CORPORATION

Volume 9, Numbers 1-2 January-February 1961

U.D.C. 534. 73 : 465. 6] : 621. 391

On Fundamental Qualities of Vowel Sounds in Speech Communication*

(Part 3)

Continued from p. 529.

***Yoshiyuki OCHIAI*†** and ***Teruo FUKUMURA*†**

4. Study of Fundamental Qualities by Further Kinds of Transmission Distortion

As seen in the preceeding chapter, the effect of distortion upon quality can be properly and correctly evaluated only by the joint consideration of its quality-destructive and quality-constructive actions. It is expected that there may be other kinds of transmission having quality effects quite different from that of BED, and thus it is very important for quality study to decide what kinds of distortion should be employed. In the case of BED we took full advantage of the distortion nature of band-cutting, by virtue of which we could distinguish the quality factor of dispersed nature from that of concentrated nature, and we could thus assure that there exists fair independency between voice and phoneme qualities. In order to find some new quality factor or to confirm the existence of the quality factors already revealed, it is necessary to further employ and study types of distortion other than BED. In this sense, quality measurement by Rotational Synchronous Distortion (RSD) was carried out in Experiments 2 and 3, and that by attenuation distortion (AD) in Experiment 6, both in parallel with that by BED. The former distortion, RSD, shifts the signal pattern upward and downward along the frequency axis and the latter, AD, shifts it downward to the audibility threshold.

4.1. Rotational Synchronous Distortion

By using different speeds during recording and reproducing on the magnetic tape recorder, it is possible to shift the pattern along the frequency axis, and we can obtain knowledge of quality factors about frequency positions from the subjective measurement carried out under this circumstance. In the physical and analytical study of chapter 1 where we designate the main formants in vowel patterns by f_1, f_2, (phonemic quality) and those in vocal patterns by F_1, F_2, (vocal quality), the two relations $f_1 : f_2 \doteqdot$ const., and $F_1 \times F_2 \doteqdot$ const. were generally observed irrespective of the variation in absolute formant positions

---

* MS in Japanese received by the Electrical Communcation Laboratory, Aug. 18, 1960. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T., Vol. 9, No. 12, pp. 1469–1549, 1960.

† Nagoya University.

due to the difference in physique and the difference in sex of the calling subjects. It is very important, from the standpoint of subjective study, to decide whether these facts observed in the physical experiment can also be confirmed or not. For this purpose we utilize RSD distortion which is thought to be the most meaningful.

4.1.1. Quality Judgment Characteristics

Fig. 20 shows the results of Experiments 2 and 3 where the percentage of correct judgment of voice and phonemic qualities are plotted for seven conditions of RSD.

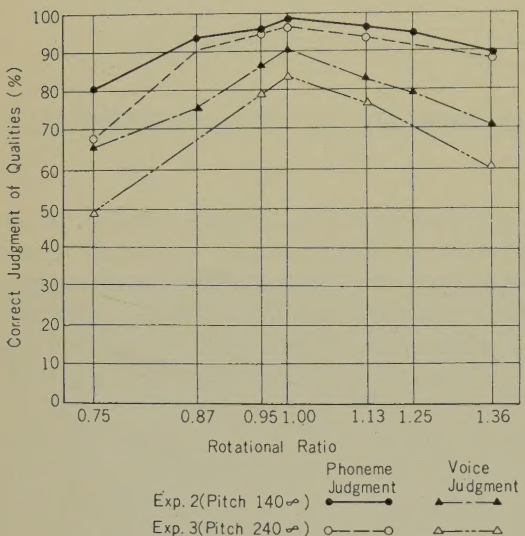


Fig. 20—Curves of voice- and phonemic-quality judgment by RSD (Experiments 2 and 3).

Both quality curves in Experiment 3 (calling pitch of 240 c/s) are lower than those in Experiment 2 (calling pitch of 140 c/s) and this may be due to the fact that the calling subjects (including males) are not accustomed to speak at the pitch of 240 c/s. The detailed study of fundamental qualities by this distortion has been published in Reference (43) and can be summed up as follows.

Phonemic qualities, which show the relation $f_1/f_2=\text{const.}$ in the analytical study, can also retain high values in subjective measurement by RSD distortion. Especially the vowels of so-called double formants. "I" and "E", which have a large ratio of f_1/f_2, are almost unaffected by this distortion and keep high quality; while the vowels of single formant, "A" and "O" are more sensitive to this distortion. (Figure is omitted.) On the contrary, those voice qualities where the relation $F_1 \times F_2=\text{const.}$ is observable, deteriorate rapidly even for small distortions (for both upper and lower shifts). Voices "K" (voice with dominance of odd-harmonics) and "Y" (voice with dominance of even-harmonics), which have many distinctive harmonics in the upper structure usually responded to this distortion most sharply. Voices "F" and "N", both having little harmonics originally and weak and poor upper structure consequently, are quite strong against this transposition distortion. From this point, it can be assumed that this distortion acts on voice qualities very differently than it acts on phonemic qualities. (Figure is omitted.) In the preceding study of BED, we revealed the fact that phonemic quality is distinctly of a concentrated nature depending on the position of concentration; that is, the formants; in other words, that the phonemic quality is definable mainly by the "frequency construction". On the other hand, by RSD we can prove that this quality is also dependent, in some measure, on the interval of f_1 and f_2. In short, phonemic quality is defined not only by the absolute location of f_1 and f_2 along the frequency axis, but also in some measure by the relative positions of f_1 and f_2, that is, the interval of the distance between f_1 and f_2. It must be noted that in Fig. 20 this dependence of the quality on the frequency interval does not hold throughout the whole range. Further detailed study on quality response will be possible only by analyzing the confusion phenomena.

4.1.2. Auditory Confusion Phenomena

In order to analyze the confusion phenomena, Table 8 is provided to present the

Table 8

CONFUSION MATRIX OF VOICE QUARITY IN RSD (EXPERIMENT 2)

r.r.: 1

	y	k	n	f	Σ
Y	258	33	7	2	300
K	18	271	9	2	300
N	12	18	250	20	300
F	1	1	13	285	300
Σ	323	289	279	309	1200

r.r.: 0. 95

	y	k	n	f	Σ
Y	267	20	13	0	300
K	31	255	12	2	300
N	11	25	252	12	300
F	2	3	15	280	300
Σ	303	311	292	294	1200

r.r.: 1. 13

	y	k	n	f	Σ
Y	228	50	22	0	300
K	35	229	34	2	300
N	15	11	253	21	300
F	0	0	13	287	300
Σ	290	278	322	310	1200

r.r.: 0. 75

	y	k	n	f	Σ
Y	182	81	29	8	300
K	63	192	32	13	300
N	33	46	196	25	300
F	9	20	42	229	300
Σ	339	287	299	275	1200

Voice: All male voices.

r.r.: 1. 36

	y	k	n	f	Σ
Y	224	44	30	2	300
K	76	187	33	4	300
N	37	29	203	31	300
F	1	11	32	256	300
Σ	271	338	298	293	1200

Table 9

CONFUSION MATRIX OF PHONEMIC QUALITY IN RSD (EXPERIMENT 2)

r.r. : 1

	i	e	a	o	u	Σ
I	239	1	0	0	0	240
E	0	238	2	0	0	240
A	0	0	230	5	5	240
O	0	0	2	223	15	240
U	0	0	0	0	240	240
Σ	239	239	234	228	260	1200

r.r. : 1.13

	i	e	a	o	u	Σ
I	234	6	0	0	0	240
E	2	238	0	0	0	240
A	0	0	239	1	0	240
O	0	1	17	204	18	240
U	0	0	3	0	237	240
Σ	236	245	259	205	255	1200

r.r. : 1.36

	i	e	a	o	u	Σ
I	216	8	0	1	15	240
E	0	238	1	0	1	240
A	0	0	233	6	1	240
O	0	0	41	174	25	240
U	0	4	2	12	222	240
Σ	216	250	277	193	264	1200

r.r. : 0.95

	i	e	a	o	u	Σ
I	238	2	0	0	0	240
E	1	239	0	0	0	240
A	0	0	218	22	0	240
O	0	0	5	208	27	240
U	1	0	0	0	239	240
Σ	240	241	223	230	266	1200

r.r. : 0.75

	i	e	a	o	u	Σ
I	237	3	0	0	0	240
E	4	188	3	9	36	240
A	0	1	154	83	2	240
O	13	7	6	141	73	240
U	4	1	4	8	223	240
Σ	258	200	167	241	334	1200

vocal confusion; that is, the confusion matrix of voice qualities, and Table 9 is for the phonemic confusion; that is, the confusion matrix of phoneme qualities. Both Tables are based on the data of Experiment 2.

In Fig. 21 is shown the confusibility between phoneme signals and phoneme qualities and in Fig. 22 is shown the confusibility between voice signals and voice qualities; both after being plotted in individual distortion conditions. With regard to the vocal confusion, the vocal patterns of individual voices are given in Fig. 23 for convenience in observing the correspondence between them.

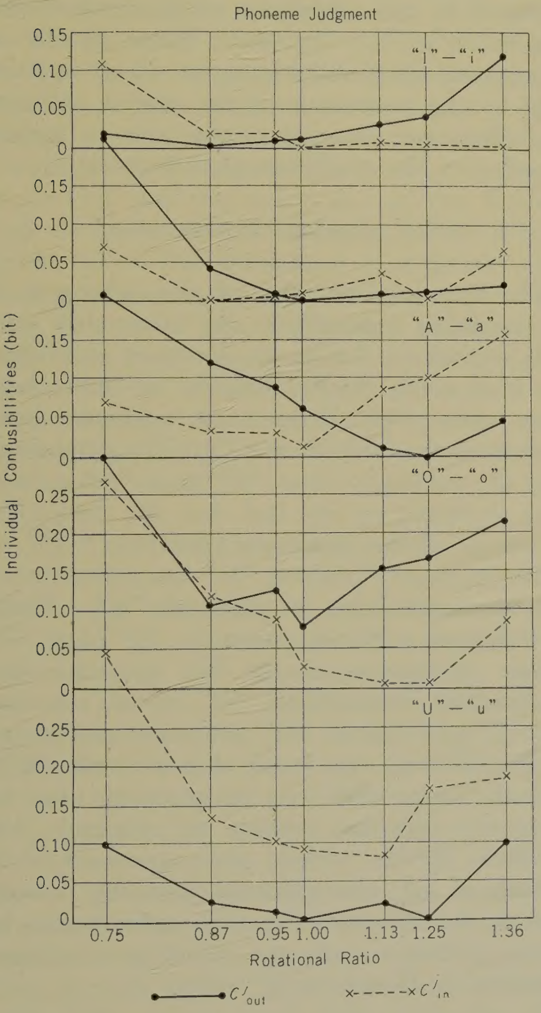


Fig. 21—Characteristics of individual confusibilities of phonemic quality in RSD (Experiment 2).

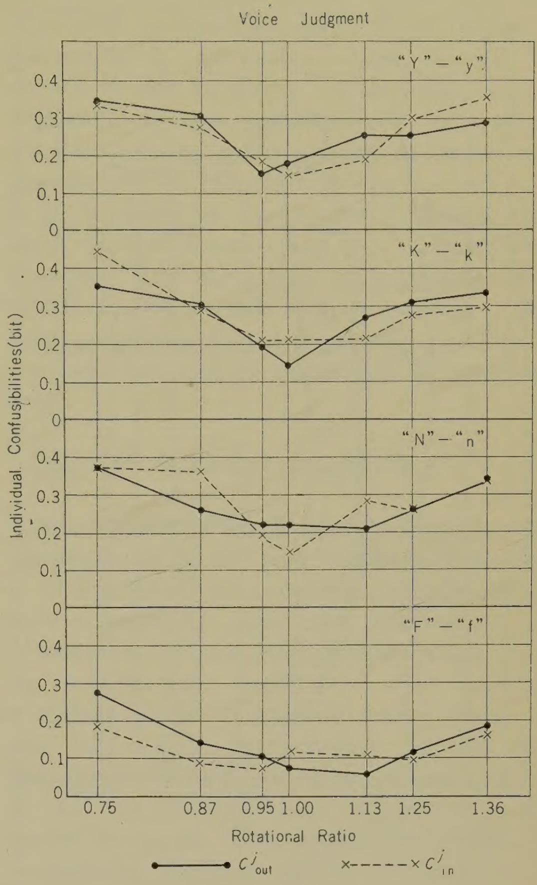


Fig. 22—Characteristics of individual confusibilities of voice quality in RSD (Experiment 2).

1) Phonemic confusion

Fig. 21 shows the individual confusion characteristics of the five phonemes in RSD, thereby stressing the distinctive difference between outgoing-(solid curve) and incoming-confusibilities (dotted curve) for almost every condition. It is easily seen that as a whole the confusion phenomena take place with some systematic regularity. These are the

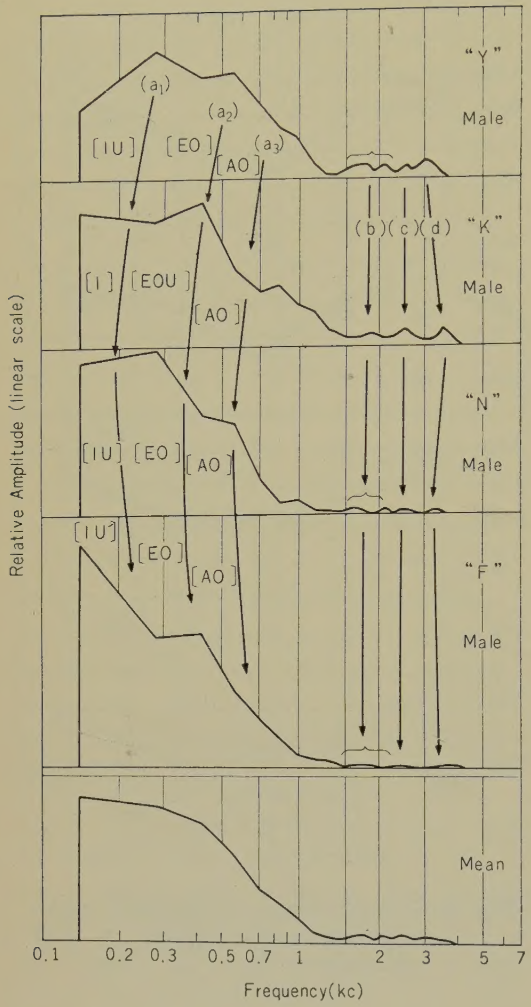


Fig. 23—Pattern arrangement corresponding to voice-quality configuration (Experiment 2).

same confusion trends that have been observed before regarding BED, and it can be said that the band concentration of phonemic factors is also detected by this distortion. Now let us examine these confusion trends for individual phonemes. For reference vowel patterns are shown in Fig. 14 (B), and in Table 9 a summary of detailed confusions is presented as the confusion matrix of five conditions of RSD.

“I”–“i” Confusion characteristic

This case shows the smallest confusibility among the five phonemes. This fact corresponds to the circumstance that this vowel has the largest phonemic double formant interval, f_1/f_2. Certain increasing of the characteristic of outgoing confusibility observed in the upper shift condition (r.r.$>$1) results clearly from the confusions outgoing from this vowel and entering into “e” and “u”, and the increasing characteristic of incoming confusibility observed in the lower shift (r.r.$<$1) is caused by the confusions coming respectively from “O”, “U”, “E”, etc. These details of confusion phenomena may be simply explained by the conparison of the relative positions of phonemic formants of the vowels participating in the confusion of this vowel.

“E”–“e” Confusion characteristic

The outgoing confusion of this phoneme is remarkable especially in the lower shift condition, where the most dominant confusion is that going chiefly toward “u” and there are also observed in some measure those which go toward “o” and “i”. The dominancy of the confusion from “E” to “u” may easily be explained by the fact that the frequency interval of the phonemic formants (f_1/f_2) of the Japanese Phoneme “U” has approximately the same value as that of “E”. Weak confusions with the phonemes “O” and “I” may correspond to the fact that the first phonemic formant of “E” comes near the position of that of “O” and sometimes of that “I”. The general trend of poor confusibility in the upper shift may be accounted for by the fact that the position of phonemic formants of this vowel never come close to those of any other vowels when the position is forced to move to the upper region by this distortion. Some incoming confusion noticed is that from “I” only.

“A”–“a” Confusion characteristic

The incoming confusion is dominant in the upper shift and the outgoing is dominant in the lower shift. This vowel may be considered as having a single formant in the phonemic sense, and its unique formant is situated in the intermediate region among the formants of all five vowels. As there is no other vowel having a single formant higher than "A", the outgoing confusion does not rapidly increase in the upper shift. As a reason why the confusibility in the condition of r.r. = 1.13 decreases more than in the reference condition it is conceivable that this vowel was probably pronounced somewhat incorrectly by some calling subject and already liable to be confused with "O" even in the reference condition, and the small distortion in the upper shift rather improved the quality of this mispronounced vowel. The trend of dominance of incoming confusion becomes most remarkable in the condition r.r. = 1.36, and the major part of the incoming confusion at this distortion point comes from "O". The dominance of outgoing confusion in the lower shift condition is exactly the reverse of the dominance of incoming confusion in the upper shift. In short, the general trend of confusion of this vowel is determined by the dominant part of it and is decided chiefly by reciprocal influences between "A" and "O".

"O"–"o" Confusion characteristic

From the reciprocity in confusion between "A" and "O", it can be easily seen that the main part of the dominant outgoing from vowel "O" in the upper shift condition is caused by confusion with neighboring vowel "A". In this shift direction, a considerable amount of outgoing confusion goes toward "u", because the second formant of vowel "O" draws near the position of that of "U". Some incoming confusion in the condition r.r. = 1.36 (though not so conspicuous) comes mainly from "U". The upper shift of the first formant of "U" gives rise to confusion with "O". In the lower shift, both the incoming and outgoing confusibilities take extremely large values at the same time. The incoming confusions in this shift direction mainly come from "A", but it is only the reversed phenomena in the upper shift and not much needs to be said. With respect to the outgoing confusions, those toward "u" prevail in this direction. This point can be explained by the consideration of relative positions of the first formant of these two vowels.

It is important to suggest the following: it is only in this vowel that two confusibilities, incoming and outgoing, take equally large values in the lower shift condition. This means that this vowel is possibly situated in the most intermediate position in the phonemic sense.

"U"–"u" Confusion characteristic

The confusion trend of this vowel is the most characteristic of the five vowels and is clearly characterized by incoming confusion dominant in both the upper and lower shifts, the two curves of both confusibilities never crossing in both shift directions. The incoming confusion in the upper shift comes from "I" and "O". The displacement of the first formant of "I" and that of the second formant of "O" by direction can account for this confusion trend. The remarkable superiority of outgoing confusion in the condition r.r. = 1.36 is mainly characterized by the dominant confusion going toward "o", and this point will be well explained by the shift of position of the first formant of "U" which results from the confusion with "O". The confusion outgoing toward "e", which is expected from the similarity of formant interval between these two vowels, is not definitely found. The reason is either that the second formant of "U" was somewhat low in position and weak in intensity, or that the distortion amount was not yet sufficient. Confusion incoming from "O" is the strongest in the lower shift, and then follows the incoming from "E". The similarity in value of f_1/f_2 explains the confusion with "E" and the neighborhood of mutual positions of the first formants explains the confusion with "O".

No conspicuous trend is found in outgoing confusion in the condition r.r. = 0.75. With regard to the formation of the phonemic quality of this vowel, it is important to notice that both the first and second formants must be taken equally into consideration for the explanation of incoming confusion. For other vowels such as "I" and "E", one of these two formants is stronger than the other in the phonemical sense. There is no vowel other than "U" which has two conspicuous formants contributing equally to the formation of phonemic qualities.

Summing up the confusion studies for the individual vowels, it may be said that all the confusion trends can be explained by the displacement of formant position by distortion shift. This means that by virture of RSD we can detect the likely kind of mechanism of phonemic quality formation as found in BED, namely, we can reveal the phonemic nature of concentration and localization of band quality and we can detect also their dependency upon frequency interval.

2) Vocal confusion

Individual characteristics of vocal confusibility for four voices with RSD distortion are shown in Fig. 22. Compared with those of phonemic confusibility, all these curves of vocal confusibility are a little complicated and there is relatively small differences between the two kind of confusibility curves. Particular attention, however, should be paid to the fact that when the distortion is still small (from r.r. = 0.87 to r.r. = 1.13), these vocal confusibility curves show more significant tendencies than those of phonemic confusibility (shown in Fig. 21) in the same domains of distortion. This is the same trend observed in the curves of general quality judgment (Fig. 20) where the general voice judgment curves show more rapid inclination of drop than those of general phoneme quality judgment in the same distortion range. This means that even a small amount of rotational synchronous distortion like this or even the small shift of pattern figures caused by the small distortion can give rise to a more serious modification in vocal quality than in phonemic quality. These vocal confusibility curves will now be examined for each combination of the voice signal and the corresponding vocal quality, and from the joint observation of the confusibility curves and the vocal pattern we want to study the essential characteristics of vocal quality element detectable by this distortion. The vocal patterns of these voices are given in Fig. 23. In this figure, the positions of corresponding pattern parts most significant in the lower frequency region are shown by the three curves with arrows, in order to facilitate the understanding of the positional relations between individual voices.

Curve "a_1" = first formant of "I" and "U".

Curve "a_2" = first formant of "E" and "O".

In the case of voice "K" only, a part of the first formant of "U" is included.

Curve "a_3" = first formant of "A" and second formant of "O", "U", and "A".

"Y"–"y" Confusion characteristic

Curve shows the characteristic form, particulary in the upper shift distortion, and shows relatively little difference between the two confusibilities in the lower shift. It is seen in the confusion matrix (Table 8) that the dominance of outgoing confusion in the condition r.r. = 1.13 is caused by the nonreciprocal confusion between "K" and "N". As explaining this trend, we can refer to the relative position of the fourth formant shown by the "d" line in Fig. 21. The contribution of this fourth formant region to the voice quality "k" and also the contribution of the third formant region given by the "c" line (probably including the fourth formant in the "d" line) to the voice quality "y" have already been detected by the BED distortion (particularly by HC 4.2 kc and 2.1 kc). Thus the importance of the formant regions in the sense of vocal quality is equally proved by the RSD distortion, the importance of which depends upon the frequency positions. By careful inspection of the lower structure, it is possible to trace that the above confusion trend also corresponds to the detail in the lower structure; that is, the confusion trend can be well explained by the relative position of the pattern parts shown by the "a_1" and

"a_2" lines. Calling subject "Y" has a smaller physique than "K" and "N", and the peak position in his pattern shown by the "a_1" line is relatively higher. By a small upper shift, it moves further to a higher position and readily comes close to the positions of the "a_2" line of "K" and "N", thus giving rise to the vocal confusions. "Y"→"k" and "Y"→"n". In the conditions r.r.=1.25 and 1.36, the "a_2" part of "K" draws near the "a_3" part of "Y" and causes an increase in incoming toward "y". The same confusion trend at these distortion points can be also explained as follows: This confusion trend is caused by the fact that the position shown by the "b" line of "K" in the upper structure draws near the position shown by the "c" line of "Y".

In the lower shift, for example in the condition r.r.=0.95, incoming confusion increases while outgoing confusion decreases. This may be explained simply by reversing what has been said for the condition r.r.=1.13 in the upper shift; that is, mutual approach of formant positions in the "d" line of "K" and "Y" as well as that of the "a_2" part of "K" and "a_1" part of "Y" can be the explanation. Sudden increase in outgoing confusion at the point r.r.=0.87 in the lower shift condition and a little dominating tendency of the outgoing confusion in the subsequent conditions can be equally explained by the consideration concerning mutual relation of formant positions in individual patterns caused by the variation of physical statures, as in the case of upper shift. In this case, however, the outgoing confusion at the point r.r.=0.75 seems to be too small, compared with the degree of dominance of incoming at r.r.=1.36, and the reason for this is not yet clear.

"K"–"k" Confusion characteristic

By rough inspection, a clear superiority of outgoing confusion in the upper shift condition and some superiority of incoming confusion in the lower shift condition are noticed, though the curve forms are in general a little complicated. In the confusion matrix (Table 8), it is seen that the conspicuous increase in outgoing confusion at r.r.=1.13 is mainly caused by an nonreciprocal confusion of this voice with "N". For an understanding of the confusion between "K" and "N", the difference in the upper structures of vocal patterns can be excluded from our consideration, because in these structures the distinct difference is found rather in intensity dimension and not in the positional factors. When we compare then the lower structures of "K" and "N", it is readily recognized that a similar systematic difference in pattern exists between them as that found between "K" and "Y". This seems to come from the difference in physiques as in the previous case: The calling subject "N" is larger in physique and thus the position of his formant is generally lower. Such a positional difference in pattern can be considered as directly corresponding to the confusion trend described above. For comparative small amount of distortion in the upper shift, the confusion coming from the voice with higher peak position ("K") and entering into the voice with lower peak position ("N") is stronger than the reverse confusion trend which comes from the voice with lower peak and enters into the voice with higher peak. This is also the case in the confusion between "Y" and "K". It seems to contradict the consideration where the positional relation found in corresponding peaks in vocal patterns in question is always and solely responsible for confusion interpretation. In such a case, it is far better to consider, as corresponding to confusion trend, the mutual positions of particular peak parts only: In this case, for example the positions indicated only by curves "a_2" and "a_3". The position shown by curve "a_3" in the "N" pattern is actually situated somewhat lower and consequently the vowel "A" of this voice has some "o"–like timbre. By this distortion, the formant of the vowel "O" (indicated as curve "a_2") of the neighboring voice "K" is shifted upward and gives rise to vowel "O" sounding with an "a"–like timbre thus resulting in vocal confusion with the voice "N". Dominance trend of the outgoing confusion at the point r.r.=1.36 is due to nonreciprocal confusion of this voice with "Y", as explained in the preceding item. In the lower shift, dominance of incoming confusion is evident, and as an explanation of it, we can utilize thc vocal pattern correspondence given by the reverse of the explanation in the upper shift distortion.

"N"–"n" Confusion characteristic

Dominance of incoming confusions at r.r.$=$1.13 and 0.87 is clear. From inspection of the confusion matrix, it can be seen that the conspicuous increase in incoming confusion in the upper shift is caused by nonreciprocal confusion of this voice with voice "K". This was already observed in the preceding item. Therefore the explanation may be almost the same as that in the preceeding item, though here is some nonreciprocity between "N" and "Y". As will be seen in the confusion matrix (not shown in Table 11), the increase in incoming confusion at r.r.$=$0.87 is due to the nonreciprocal confusions between "N" and "Y" and between "N" and "F". These trends can also be considered as corresponding to the mutual relation of peak positions in the vocal patterns concerned. For nonreciprocity "N"→"k" which occurs in the lower shift condition and which contradicts the validity of the uniform positional relation found in every peak structure, special attention should also be paid here and the special positions of particular peak structures should be stressed primarily, as stated before.

"F"–"f" Confusion characteristic

Incoming confusion is dominant at r.r.$=$1.13 and outgoing confusion is generally dominant in the lower shift. But the dominance of incoming confusion in the upper shift is not worthy of mention, because it is almost the same as in the reference condition. It is seen from the confusion matrix that dominance of outgoing confusion in the lower shift is caused by nonreciprocal confusion of this voice with "N" and with "K". The mutual relation of peak positions in pattern structure is also a reasonable explanation of these characteristic confusions of this voice with the neighboring voices having comparatively strong fundamentals.

Reviewing all the above considerations, it is generally concluded that most of the nonreciprocal confusions taken into consideration are explained by the positional relationship found in the corresponding peak structures of the vocal patterns concerned. It is accordingly possible to say that the RSD distortion can reveal the delicate element of voicequality contributed by the position and sharpness of the formants; in other words, we can obtain the importances of the peak and glenstructure of the vocal pattern which defermine the real situation of the so-called formants. It is a unique merit of this distortion that such fine structure of vocal patterns as vocal nuance can be detected and there is no other distortions serviceable for revealing such a fine structure.

4.2. Attenuation Distortion

If the quality distribution in the *frequency domain* can be obtained by quality measurements of the BED distortion, quality measurements should be made of the AD distortion to obtain the quality distribution in the *level domain.* Differing from the BED distortion where the sensation area is cut in the vertical direction only, this AD distortion brings gradual change of pattern caused by cutting parallel to the audibility threshold. Further, the meaning of the peak part of patterns in the quality view may be clarified by analysis of the quality response to the signal stimuli brought near the threshold. Apart from the pure theoretical viewpoint of speech quality, subjective measurements in AD is also necessary for the rating of speech communication systems. However, the problem of system rating is not treated here.

4.2.1. Quality Response Characteristic

Mean characteristic curves obtained by averaging measurements of the two qualities, phonemic and vocal, at several distortion conditions of AD are shown in Fig. 24. The solid curve stands for the phoneme quality characteristic, and the dotted curve the voice quality characteristic.

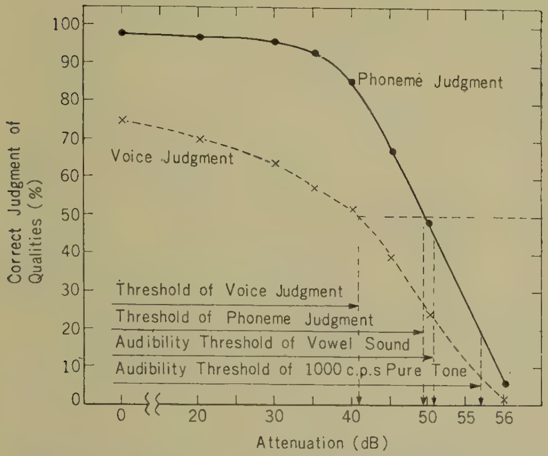


Fig. 24—Characteristic curves of phonemic and vocal qualities in AD (Experiment 6).

After retaining high percentages in small distortion conditions, the phoneme quality curve shows a sudden fall which begins from the attenuation of about 35 dB. The voice quality curve, on the contrary, falls from the begining with a gentle slope, and falls always with varying slope. Such a remarkable difference between the two qualities was already recognized also in BED. The quality response which behaves not too sharply but gently against both the band and level distortions seems to belong to the voice quality. The slope of the voice quality curve becomes fairly large in the neighborhood of 40 dB attenuation and this point coincides approximately with the begining of the elision phenomena. In Fig. 24 are indicated further values of various kinds of thresholds. The threshold of 50 % audibility for a 1,000 c/s pure tone has the lowest value, corresponding to the attenuation of about 57 bB. At this point, the percentage of the phonemic quality is about 27 %, and that of the voice quality judgment is about 5 %. The mean threshold* of 50 % audibility averaged for all vowel signals corresponds to the attenuation of about 51 dB.

As corresponding to the 50 % correct judgment in phonemic quality; that is, as the threshold of phonemic judgment, attenuation of about 49 dB is determined. This value is not so very different from that of the audibility threshold of vowels (51 dB), but from this superficial fact, it should not be immediately concluded that most of the audible vowel signals have been correctly judged with regard to phonemic quality. The reason, as will be explained later in more detail, is that there is only 34 % (not 51 %) elision in the phoneme judgment at this attenuation point. This fact seems to reflect that the listeners paid more atention when they are forced to judge the phonemic quality of signals than when they are simply forced to judge the audibility of signals. Likewise the attenuation of about 41 dB can be obtained as corresponding to the 50 % threshold of the voice quality. It is interesting and even important to notice that this threshold attenuation for the voice quality is much less than

Table 10

PHONEME CONFUSION MATRIX IN AD (EXPERIMENT 6)

Att. 30 dB

	u	i	e	o	a	el.	Σ
U	391	7	0	2	0	0	400
I	3	396	1	0	0	0	400
E	2	1	397	0	0	0	400
O	5	0	3	351	41	0	400
A	1	0	0	14	385	0	400
Σ	402	404	401	367	426	0	2000

---

* This threshold was obtained by the listeners for the particular criteria of binary judgment, "audible" or "inaudible," instead of making ordinary quality judgments.

that for the phonemic quality.

Table 10

Att. 40 dB

	u	i	e	o	a	el.	Σ
U	322	30	4	4	3	37	400
I	29	338	5	0	1	27	400
E	12	1	378	4	0	4	400
O	15	1	21	331	25	7	400
A	6	0	1	42	348	3	400
Σ	384	371	409	381	377	78	2000

Att. 50 dB

	u	i	e	o	a	el.	Σ
U	117	31	22	8	13	208	400
I	16	145	22	3	12	203	400
E	18	18	244	8	4	108	400
O	22	7	45	180	35	111	400
A	11	2	7	44	271	65	400
Σ	184	203	340	243	335	695	2000

4.2.2. Confusion Phenomena in AD

1) Phonemic confusion

The confusion matrices for the attenuations of 30, 40, and 50 dB are shown in Table 10. In this table are arranged vowels and phonemes so that confusion trend at low level may easily be found.

The blocks outlined with solid lines show domains of grouping most conspicuously recongnized, while those outlined with dotted lines show not so distinct ones. At the attenuation of 30 dB, the group between "A" and "O" are most distinct and the group between "U" and "I" is also recognizable. Though it is without much clearness, there is another group composed of "U"–"I"–"E"–"O," revealing thereby a confusion trend coming from individual vowels and going to "u." These trends of composing group increase more remarkably from the attenuation of 40 dB and the groupings between "U" and "I" and between "A" and "O" become far more distinct. The grouping between "E" and "O" is also conspicuous but a somewhat nonreciprocal, and the grouping in "U–I–E–O" also becomes distinct, thereby indicating the confusion coming from individual vowels and going to quality "u." From the joint consideration of the vowel patterns compared with the confusion characteristics, it is noticed as common trends that there exist distinctive formants having close mutual relations in the individual patterns which compose a certain group. The reason why the nonreciprocal confusion happens in the group of "E–O" may be explained as follows: 1) the frequency positions of the first formants of these two vowels are close to each other; 2) "O" has only a single formant and its unique quality factor seems easily to disappear at the lower level, while for "E" the second formant is a far more important factor than the first; further the second quality factor seems to remain, still audible even at the lower level. It may be suggested by the same nature of things that the phonemic circumstance of plural formants causes "E" being confused more easily with "U" rather than with "O," because the first formant of "E" is situated closer to the formant of "O" than that of "U."

The elision phenomenon increases at the attenuation of 50 dB, and the confusion trends at the attenuation of 40 dB begin to be of an ambiguous nature, with some traces still remaining to a certain extent.

In summing up, the contents and details

of the phoneme confusion in AD may reasonably be explained by examining whether the phonemic formants remain audible or not in the neighborhood of the audibility threshold. Further we can say that the manner in the phoneme judgment in the lower level seems depend upon the recognizability of peak positions in vowel patterns brought down-ward to the threshold. If the peak in question is really an important part, its phonema can still be recognized even when that peak approaches extremly close to the boundary of the audible domain. And, as soon as the peak goes down below the threshold, the recognition of phoneme is suddenly destroyed and consequently the percentage of correct phoneme judgment abruptly diminishes. These statements tell us nothing but the band-localizability of quality concentration and band-independency of quality recognition for individual phpnemes. It is also noteworthy that the conspicuous quality-formative action is never observed in this AD distortion throughout all the distortion ranges. AD distortion does not present so many things as BED, so long as the mechanism of the formation of phonemic quality concerns.

2) Vocal confusion

The representative confusion matrices at the attenuations of 30, 40, and 50 dB are shown in Table 11.

The manner of showing the confusion groupings by block is the same as in the previous phonemic case. The confusion matrices in BED are also given in Table 12 for the convenience of comparison, and the matrix of this voice group in the reference condition is shown in Table 2. It is readily seen that the confusion trend observed in the reference condition is still conspicuous even at the attenuatjon of 30 dB; further, in this stage of distortion the trend of non-reciprocal confusion is also entirely the same, with only the difference that thereby the voice judgment in general becomes a little ambiguous. From the attenuation of 40 dB the grouping of "M–H" becomes less conspicuous and the outgoing confusion from

Table 11

VOCAL CONFUSION MATRIX IN AD
(EXPERIMENT 6)

Att. 30 dB

	t	s′	m′	h	el.	Σ
T	295	137	49	19	0	500
S′	102	284	105	9	0	500
M′	44	87	303	66	0	500
H	12	37	46	405	0	500
Σ	453	515	503	499	0	2000

Att. 40 dB

	t	s′	m′	h	el.	Σ
T	281	125	58	35	3	500
S′	105	229	111	43	8	500
M′	32	109	238	86	35	500
H	33	52	68	288	57	500
Σ	455	515	452	452	103	2000

Att, 50 dB

	t	s′	m′	h	el.	Σ
T	158	83	58	55	146	500
S′	77	147	83	28	165	500
M′	28	52	97	32	291	500
H	32	57	47	80	284	500
Σ	295	339	285	195	886	2000

Table 12

VOCAL CONFUSION MATRIX IN BED (EXPERIMENT 6)

HC : 1.5 kc

	t	s′	m′	h	el.	Σ
T	203	201	68	26	2	500
S′	78	234	155	31	2	500
M′	76	107	269	46	2	500
H	66	46	49	339	0	500
Σ	423	588	541	442	6	2000

LC : 1.0 kc

	t	s′	m′	h	Σ
T	284	154	53	9	500
S′	144	258	91	7	500
M′	42	110	277	71	500
H	7	17	33	443	500
Σ	477	539	454	530	2000

HC : 0.7 kc

	t	s′	m′	h	el.	Σ
T	121	192	94	77	16	500
S′	98	172	157	66	12	500
M′	86	158	162	77	17	500
H	94	91	85	218	12	500
Σ	394	613	498	438	57	2000

LC : 2.5 kc

	t	s′	m′	h	Σ
T	211	179	91	19	500
S′	136	224	116	24	500
M′	60	102	213	125	500
H	7	17	38	438	500
Σ	414	522	458	606	2000

"H" comes to lose its particular destination (object of confusion). This trend conforms entirely to that observed at HC 1.5 kc in BED and this means that the disappearance of higher noise components from the voice "H" is taken in common consideration as the cause of this trend. From the attennation of 50 dB all the groupings are broken down, but only the group of "S′–M′" retains its trace. This resembles the confusion trend brought about by HC 0.7 kc in BED, although with the less equivocation in the HC case. As a whole distortion AD effects the vocal confusion in almost the same manner as in HC distortion but the former is less characteristic. This may be explained as follows: As the vocal patterns originally have envelopes similar to those caused by the frequency characteristic of a low pass filter with gradual attenuation, the AD distortion, roughly speaking, brings about a very slow high frequency cutoff effect of pattern figure without such marked emphasis on suppression of a particular band accompanying as is usually seen in the case of BED.

After all, the AD distortion is not conceived as the decisive distortion of interest, while RSD is a very interesting distortion which furnishes new knowledge on voice-quality construction. In the latter case, it became clear that voice quality depends not only upon the outline of envelopes of vocal patterns, but also on its detailed construction; in other words, on the delicate positional information of formants in vowel patterns. Further study should be made as to how such voice-quality factors are inter-related to phonemic-quality factors. As for the phonemic quality, its quality factors are much more band localized and depend strongly upon the frequency position. This was detected equally in the experiments by the RSD and AD distortions, the former also revealing the dependency of this quality upon the frequency interval of constitutent factors. Under these circumstances, there is possibility to introduce a more proper scheme of pattern expression to represent the phonemic quality. The mechanism of quality formation of "U" is the most complicated among the five vowels and the representation of its timbre pattern will be the most difficult one. This is probably due to the characteristics of "U" in Japanese.

Discussion and Summary

The word "quality" as used in "speech quality" or "transmission quality" has a very important concept in speech communication or speech transmission. The evaluation of a subjective system a system composed of a combination of caller and listener) will be possible by introduction of an objective transmission system and by application of transmission techniques, meanwhile the rating of a physical transmission system can be made solely by the response of the subjective system (namely, speech quality). The evaluation of a subjective system and the rating of an objective system are in a relation of interdependence: According as the one is clarified, the other will be brought into clearness by itself.

In the present study the first effort has been made solely by centering on the subject of vowel timbres in order to introduce and define a new concept of quality called "voice quality", although our study was originally started from the concrete and practical problem of the rating of speech transmission systems. To clarify "voice quality" it is necessary to put it in contrast with "phonemic quality" which was the only concept of the traditional articulation transmission, and thus the monism of quality in the past should be replaced by the dualism. The "phonemic quality" and "voice quality" are embodied inseparably in an actual structure of vowel timbres. Overcoming this difficulty of indivisibility between "voice" and "phoneme" in vowel sounds, each of the two qualities was extracted by the proper methods, and as the first stage of our study it was proved that these two qualities, phonemic and vocal, are quite different from each other and that the one is neither subordinated to the other nor predominant over the other. Of course, because of the fundamental nature of the qualities, one can not entirely be independent of the other; they are probably most intimately connected with each other by some law which is not yet known. Be that as it may it is however very necessary to insist that these two qualities can have their own proper "raisons d'étre" and therefore it is possible to prove that they differ from each other. The methods adopted here can be divided into the following three. The first is the method of deriving two kinds of vowel timbre patterns corresponding to the two quality concepts, phonemic and vacal. They are called "phonemic patterns" and "vocal patterns". The second is the actual subjective measurement of the timbre quality by impressing the timbre signals corresponding to the respective patterns as our objective signals. It means actual observation by subjective measurements of the values of the two qualities. This observation has been continued for many years in our Laboratory using various combinations of voice variaties and listener groups. To make this kind of experiment very effective, only observation under distortionless condition is not sufficient and thus the observation must be carried out on speech qualities in various conditions where transmission distortions are introduced. We can study by this method the exact relation or "correspondence" between "deformed patterns" and "modified qualities" and this serves for the study of quality interpretation designated as the third item of our study program.

To clarify and reveal the meaning of the quality construction contained in the timbre patterns, we used chiefly the BED distortion, because timbre patterns are usually considered as "frequency" construction or "harmonics" construction. This distortion is a combination of systematic, directional and at the same time gradually increasing high-cutting and low-cutting. In addition to this, RSD and AD distortion were also used. It is clearly pointed out that the phonemic and vocal qualities respond characteristically and differently for the application of distortions: for distortion where either the band is cut off gradually from upper band to lower or from lower band to upper; or where signal patterns as a whole are gradually shifted to a higher position or to a lower position; or where finally signals are gradually lowered as a whole toward the auditory threshold. Concerning the point that the responses of phonemic and vocal qualities differ of course depending upon the type of distortion, we can say that it was a sufficient to prove that these qualities were different from each other. The inner mechanism of quality formation, for example phonemic and vocal qualities, is not yet clarified. To complete our quality study, the study of auditory timbre confusion should be considered for minute interpretation of quality phenomena.

The quality is diminished by introduction of transmission distortions, because the qualities are destructed by distortions. Therefore precise study of quality deterioration should be based upon the minute observation of quality destruction mechanism. As the second becomes lower, mishearing is more likely to occur. As the sound is destroyed, ambiguity increases and mishearing occurs. Study of mishearing, that is, of confusion will be a principal theme for minute and precise quality study. It is a study of greater significance and of more attraction than the analysis of correct hearing phenomena. However, at this stage, attention should be paid to the fact that a study of confusion, requires voluminous data, troublesome to arrange; and further that it may be temporarily sufficient from the practical standpoint of the problem but insufficient from the standpoint of timbre quality theory, when the contents of confusion are based upon the details of quality deterioration caused by so-called mishearing. Significance of confusion in quality theory will be enhanced by considering not only outgoing confusion but also incoming confusion, as the former clarifies the mechanism of quality destruction and the latter that of quality formation. Introduction of the concepts of "outgoing" and "incoming" confusion bring double meanings in quality of speech, giving respectively "+" and "−" meanings to formant theory. As this point is exceedingly important for quality theory, phonemic quality was considered for more concrete explanation.

As a condition governing phonemic quality such and such formant has been said to be placed in such and such frequency position in the structure. This may surely be a necessary condition to define phonemic quality, but it is not sufficient. To make the definition perfect, it should be added that dominance should not be placed in such and such positions. Then necessary and sufficient definitions will be those which describe both positive and negative importances as follows.

(1) Dominant part in structures should be placed in such and such positions: description for peak structure with regard to patterns, and proposals of "positive importance" in respect of qualities.

(2) Dominant part should not be placed in such and such positions in the structure: description for glen structure with respect to patterns, and proposals of "negative importance" with regard to qualities.

As a definition of speech quality, these two conditions must be fully considered. Consideration of confusion, as an advanced step in our study of internal phenomena of qualities, yet limited to confusion as quality destruction phenomena due to distortion, so long as it considers confusion among articulation loss.

(This is outgoing confusion). Quality destruction phenomena, though it may seem strange at first glance, allow the natural explanation that incoming confusion should be considered by seeking quality momentum forming in the concept of frequency band.

Voice quality may be viewed from the same standpoint, too. The distribution of voice qualities, different from phonemic qualities is rather dispersive and non-concentrated as already explained. In other words, for this quality concrete and fine forms of envelope outline of vocal patterns are also responsible. To phonemic quality, it is "positive importance" by peak structure that plays a leading part in contribution. In this case "negative importance" by glen structure has only a subsidiary and supplementary meaning. In voice quality, positive and negative importances are quite comparable in their influence and quite concurrent in their action. It is necessary to add a notion that the peaks and glens in structures must be considered in mutual relation between all signal patterns and must not be evaluated by the absolute values solely in the pattern of the signals considered. Here we try therefore to compare the individual patterns with an averaged pattern in order to bring the "peaks and glens" in bold relief. Comparison with an averaged pattern is in some measure useful in the consideration of quality order in the signal ensemble. By the comparison process of phonemic patterns with phonemic qualities and of vocal patterns with voice qualities, it was first possible to concretely describe and define the respective qualities. This shows that these patterns have been appropriately introduced.

Viewed from a physical standpoint, wave forms or pattern types are simply as such, expressing nothing else, notregardless of the whether waves and patterns are expressed on the time axis or frequency axis. When we are inclined to add some meaning or importance to them, mistakes are likely to happen, as clearly ahown by the past mistakes in formant theory. For example, in physical aspect, it is nonsense to determine a *hub* or *navel* at some point in waves or patterns. When we reflect to quality, the matter is different. It is because sensation is limited, and in the limited sensation there is some domain of so-called "adaptation". Then there should be the center in the region of adaptation and of course there may be the regions of "extension". It is not mistaken but rather reasonable to expect the centers of qualities in the physical patterns in association with the principal correspondence to quality formation, or to evaluate the significances of pattern structure as corresponding to importances in the mechanism of quality formation. The introduction of a phonemic or vocal formant, if mechanically named for the physical structure of the patterns, may be irrational and dogmatic, but it is perfectly rational and reasonable if introduced after careful examination by quality consideration.

The quality-balancing point is used here as one of the central concepts. It may be one of the clues serviceable in quality interpretation. Transmission Sience in the past could have the concept of "bandwidth", but could not have that of band-center. Band-quality as well as level-quality should have some concept of "hub" or center. It is very strange that there were only concepts of frequency band and dynamic range, but no concept of "center". The base in transmission science with the center lacking must be refounded.

It is very natural that the central concepts are different, as they may be, according as the quality differs. For example, the center of voice quality comes from the fact that it is divided into higher and lower bands and widely dispersed and spread, bordering vocal glen which appears clearly in the system of vowels of Japanese. In phonemic quality there are two distinct centers, one being the center of "+" importance and the other of "−" importance. Individual phonemes have quality accumulative points as belonging to their own phonemic quality, which becomes the center of "+" importance. They also have "−" quality band of their own, in the midst of which exists the center of "−". The crossing point of both characteristic curves by LC and HC was pointed out as appearing generally in the center of or near the distribution center of phonemic quality. More precisely considered, it is with reason that the crossing point of incoming confusion generally conforms with this "+" quality accumulating point and that the crossing point of outgoing confusion coincides with the center of this "−" quality band.

References

(1) H. K. Dunn: Calculation of vowel reasonances, and an electrical vocal tract, *J.A.S.A.*, **22,** 6, Nov., 1950.

(2) C. G. M. Fant: On the predictability of formant levels and spectrum envelopes from formant frequencies, Reprint from "For Roman Jakobson", 1956.

(3) G. Fant: Acoustic Analysis and Synthesis of Speech with Applications to Swedish. Ericsson Technics. No. 1, 1959.

(4) Y. Ochiai, T. Yamashita and T. Fukumura: On transmission qnality, *Res. Rep. Fac. Eng., Nagoya Univ.* Ⅲ–1, 1950. (in Japanese)

(5) Y. Ochiai and T. Yamashita: Logic of fidelity, *Res. Rep. Fac. Eng., Nagoya Univ.*, Ⅲ–2, 1950. (in Japanese)

(6) Y. Ochiai: On naturalness, *Res. Rep. Fac. Eng., Nagoya Univ.*, Ⅳ–1, 1951. (in Japanese)

(7) H. Fletcher and R. H. Galt: The perception of speech and its relation to telephony, *J.A.S.A.*, **22,** 2, March, 1956.

(8) B. S. Cohen: Apparatus standards of telephone transmission, and the technique of testing microphones and receivers, *I.E.E. Journal*, **66,** 374, Feb., 1928.

(9) K. Braun: Die Bezugsdämpfung und ihre Berechnung aus der Restdämpfungskurve eines Übertragungsystem, *T.F.T.*, Aug., 1939.

(10) K. Braun: Theoretische und experimentelle Untersuchung der Bezugsdämpfung und der Lautstärke, *T.F.T.*, **29,** 2, Feb., 1940.

(11) K. Braun: Eine neuere Bezugsdämpfungsmesser mit objektiver Erregung und Anzeige, *T.F.T.*, **29,** 8, Aug., 1940.

(12) P. W. Blye, O. H. Coolidge and H. R. Huntley: A revised telephone transmission rating plan, *B.S.T.J.*, XXXIV, **3,** May, 1955.

(13) H. Schulz: Die Fernsprechbetriebsmasse. Ihre Entstehung, Bedeutung, kritische Würdigung und Verallgemeinerung, *E.F.D.*, Okt., 1934.

(14) F. W. McKown and J. W. Emling: A system of effective transmission data for rating telephone circuits, *B.S.T.J.* July 1933.

(15) K. Braun: Die Bedeutung und Bestimmung der Übertragungsgüte in Fernsprechverkehr, *T.F.T.* **29,** 5, Mai, 1940.

(16) N. R. French and J. C. Steinberg: Factors governig the intelligibility of speech sounds, *J.A.S.A.*, **19,** 1, Jan., 1947.

(17) T. Miura: Quantitative relation between transmission quality and transmission characteristics, Electr. Commun. Labor. Techn. J. NTT. Jap., 3, 4, p. 468, 1954. (in Japanese)

(18) N. B. Pokrovskiy, A. F. Beletskiy, E. Y. Gurbanov i dr: Raschet i izmereniye razborchivosti ruskoy rechi, *Tr. VKIAS.* 33, 3–159, 1952.

(19) E. Yu. Gurbanov: O prenenenii razlichnykh vidov artikulyatsii dlya eksperimentalynogo opredeleniya razborchivosti peredachi rechi cherez telefonnye i radiotelefonnye trakty, *Tr. VKIAS.* 35, 3–40, 1953.

(20) A. A. Tkachenko: Tonalnyy metod izmereniya razborchlvosti rechi, peredavayemoy traktami svyazi, *Akust. Zh.* Ⅰ, **2,** 171–180, 1955.

(21) Yu. S. Bykov: Teoriya razborchivosti rechi vliniyakh svyazi, *OBORONGIZ*, 1954.

(22) M. A. Sapozhkov: Ob odnoy osobennosti razborchivosti formant, *Akust, Zh,* Ⅴ, **2,** 212–214, 1959.

(23) J. Collard: A theoretical study of the articulation and intelligibility of a telephone circuit, *Elect. Comm.*, Jan., 1929.

(24) J. Collard: A new criterion of circuit performance, *Elect. Comm.*, April, 1933.

(25) J. Collard: The practical application of the new unit of circuit performance, *Elect. Comm.*, April, 1934.

(26) F. Lüschen und K. Küpfmüller: Die Entwicklung der Übertragungstechnik für den Nachrichtendienst uber Leitungen, Jahrbuch der elektrischen Fernmeldewesens von F. Gladenbeck, Berlin, 1938.

(27) E. Schäfer: Uber die Hörbarkeit von Frequenzbandänderungen bei der Ubertragung von Sprache, *E.N.T.*, **15,** 8, Aug., 1938.

(28) P. R. Arendt: Gedanken zur Bestimmung der Übertragungsgüte. *E.F.D.*, 50, p. 326–332, Okt., 1938.

(29) D. L. Richard and J. Swaffild: Assessment of speech communication links, *P.I.E.E.*, 106, 1959.

(30) Y. Ochiai: Study on naturalness of speech, *Progress Report of Electr. Commun. Labor.*, NTT, No. 644, Feb., 1955. (in Japanese)

(31) Y. Ochiai: Study on naturalness of speech, *Progress Report of Electr. Commun. Labor.*, NTT, No. 809, Feb., 1955. (in Japanese)

(32) Y. Ochiai: Study on naturalness of speech

(II), study on vocal and phonemic confusions, *Progress Report of Electr. Commun. Labor.*, NTT, Feb.. 1956. (in Japanese)

(33) Y. Ochiai: Study on naturalness of Japanese vowel, *Progress Report of Electr. Commun., Labor.*, NTT, No. 87, Mar., 1957. (in Japanese)

(34) Y. Ochiai: Study on the problems of representation and measure of naturalness of speech, *Progress Report of Electr. Commun. Labor.*, NTT, No. 118, Mar., 1958. (in Japanese)

(35) Y. Ochiai: Transmission of quality, *M.F.E. Nagoya Univ.*, 6, 2, Nov., 1954.

(36) Y. Ochiai and T. Yamashita: On timbre quality, *M.F.E. Nagoya Univ.*, 7, 1, May, 1955.

(37) Y. Ochiai and T. Fukumura: Timbre study of vocalic voices, *M.F.E. Nagoya Univ.*, 5, 2, Sept., 1953.

(38) Y. Ochiai and T. Fukumura: Timbre study of vocalic voices viewed from subjective phonal aspect. Part I, Prelimiuary studies on naturalness and articulation qualities actually and directly measured with respect to band-eliminating distortion, *M.F.E. Nagoya Univ.*, 8, 1, May, 1956; Part II, Preliminary studies on timbre confusion of phoneme and voice, Part III, Generalized treatment of timbre confusion, *M.F.E. Nagoya Univ.*, 8, 2, Nov., 1956,

(39) Y. Ochiai and T. Fukumura: Preliminaries to analysis of quality in speech communication, *M.F.E. Nagoya Univ.*, **9**, 2, Nov., 1957.

(40) Y. Ochiai and T. Fukumura: Beiträge zur Erkenntnis der Klangfarbestruktur bei vokalischen Klangbildern, *M.F.E. Nagoya Univ.*, 8, 1, May, 1956.

(41) Y. Ochiai and T. Fukumura: Some views from quality phenomena, *Paper read at the annual meeting of the Acous. Soc. Jap.*, May, 1956.

(42) Y. Ochiai and T. Fukumura: On phonemic confusion. *Paper read at the annual meeting of the Acous. Soc. Jap*', May, 1955.

(43) Y. Ochiai and T. Fukumura: On vocal confusion, *Paper read at the annual meeting of the Acous. Soc. Jap.*, May, 1956.

(44) Y. Ochiai and T. Fukumura: Study of fundamental qualitids of vocalic timbre by rotational synchronous distortion, *M.F.E. Nagoya Univ.*, 8, 1, May, 1956.

(45) R. Husson; Acoustique et physiologie phonatoires, *Journal de Physique et le Radium* (*Physique Appliquée*) *T.* 18, Mars 1957.

(46) R. Husson et L. Pimonow: Facteurs acoustiques des voyelles humaines (non nasalisées) et fréquence de conpure du pavillon pharyngobuccal, *Comptes Rendus des Séances de 1' Académie des Sciences.* T. 244, Février 1957.

(47) R. Husson: Einige Ergebnisse von Untersuchungen der akustischen Struktur der Vokale. *Phonetica.* Vol. 2, No. 3/4, 1958.

(48) R. Husson: Physiologie phonatoire du larynx. (Anatomie et Physiologie du Larynx.) *Oto-Rhino-Laryngologie.* 11, 1955.

(49) A. Moles: Sur la "caractérisation" du discours et de la diction. *Cahier d'Acoustique.* No. 76, 1957.

* * * *

Review of the Electrical Communication Laboratory
Volume 9, Numbers 1-2, January-February, 1961

U.D.C. 621.375.9 : 621.375.121.1

Staggered Operation of Doubly Resonant Parametric Amplifiers*

Shigetaka HAMADA† and *Hisakazu MUKAI*†

To obtain a broad bandwidth, a staggered operation of doubly-resonant variable-capacitance parametric amplifiers in the 6 Gc band has been studied. A doubly-resonant amplifier has been achieved by inserting a resonant stub into a waveguide cavity. The experimental results, especially the improvement of bandwidth characteristics by adjusting the phase difference between the pumps of two parametric amplifiers connected in cascade are reported. For a gain of 29 dB, a bandwidth of 55 Mc has been obtained.

Introduction

Since the invention of the parametric amplifier, many resonant amplifiers have been designed. But they all suffer from the limitation that the bandwidth, which is determined by the Q of the turned circuit, is narrow. As a method to improve the bandwidth characteristic, staggered operation of several resonant-circuit parametric amplifiers is considered.

In a resonant-circuit parametric amplifier for staggered operation, the resonant-circuit must be tuned to both the signal and the idler and the pump frequencies. An amplifier of this type, i.e. a doubly-resonant parametric amplifier which may be conveniently connected in cascade has been designed. This amplifier consists of a resonant waveguide cavity and an added resonant element, and can be tuned to two frequencies with an arbitrary frequency difference.

Let two such amplifiers be connected in cascade through a circulator. As both the output signal and the idler of the first amplifier enter the second amplifier, the performance of the second amplifier is different from that of the first amplifier, to which no idler is fed. This effect will be useful for improving the resultant gain characteristic of resonant-circuit amplifiers.

1. Doubly-Resonant Parametric Amplifier

1.1. Construction

The construction of the doubly-resonant parametric amplifier is shown in Fig. 1. The signal is fed into the amplifier through the waveguide *a* shown in Fig. 1, and the ampli-

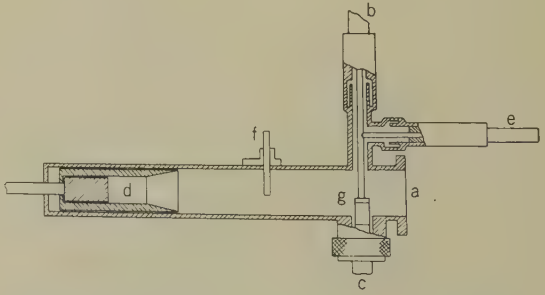


Fig. 1—Design features of the doubly-resonant parametric amplifier.

---

* MS in Japanese received by the Electrical Communication Laboratory on 27 May, 1960. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T., Vol. 9, No. 10, Oct. pp. 1193–1200, 1960.

† Electronic Parts Research Section.

fied signal and the idler also flow out through *a*. Being of the non-directional resonant type, this amplifier needs a circuit such as a circulator or a directional coupler to separate the output signal from the input signal. Pump power is supplied to the semiconductor diode *g* through the coaxial line *b*, while d.c. bias is supplied through *c*. The waveguide cavity between *g* and the short piston *d* forms a resonator for about 6 Gc, and a stub *f* is inserted into it as another resonant element to make a doubly-resonant circuit. The short piston has a double-stage construction to have a different effect on the signal and pump. The circuit to adjust the susceptance of the crystal and its supporting circuit is shown at *e*. It, together with the crystal and its supporting circuit, can also be used as a resonant element in place of the stub to obtain a doubly-resonant circuit.

1.2. Principle of Operation of the Amplifier

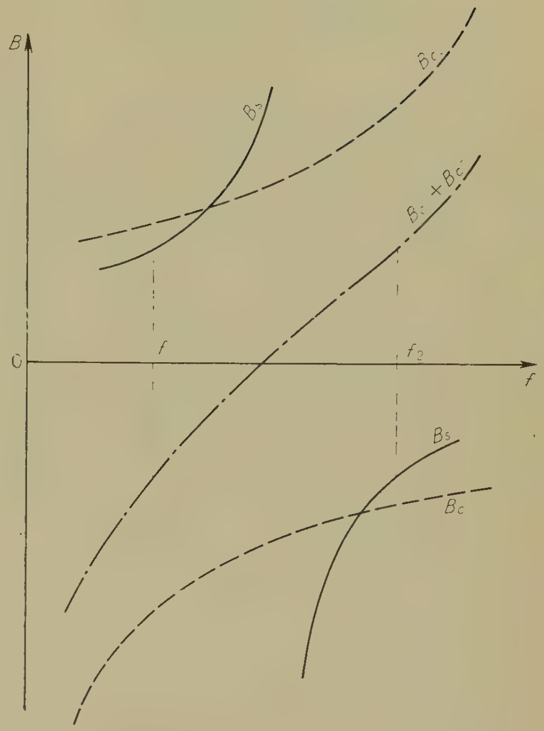
There are two methods to get a boubly-resonant circuit in this amplifier; one with a stub as an added resonant element, and the other with the crystal circuit itself.

1) The case using stub resonance.

The stub is inserted near but slightly apart from a standing-wave node in the cavity resonator. Then the susceptances of the cavity at the stub position, looking towards the short piston and looking towards the semiconductor diode are as shown by the dotted curves of Fig. 2. On the other hand, the susceptance of the stub is shown as a solid curve. The summation of these susceptances becomes zero at two frequencies, f_1 and f_2.

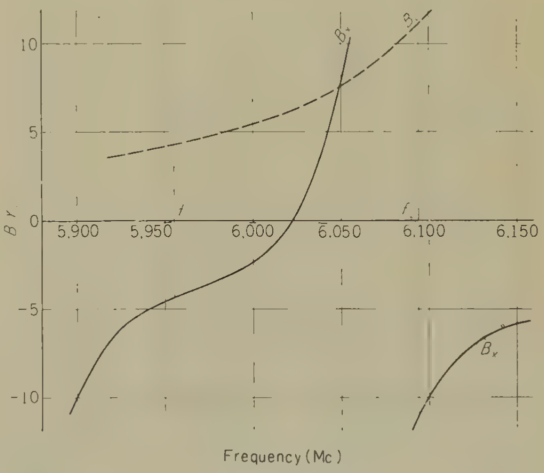
2) The case using the resonance of diode circuit itself.

The susceptance of the diode circult, which consists of the diode, its supporting coaxial line, and a coaxial short piston, has the characteristic shown by the solid curve of Fig. 3. In the figure the dotted curve shows the susceptance of the cavity looking towards the short piston. The summation of these susceptances again becomes zero at two frequencies, f_1 and f_2.



B_{c1}: Susceptance of the cavity at the stubpoint looking towards the short piston.
B_{c2}: Susceptance of the cavity at the stub point looking towards the semiconductor diode.
B_s : Susceptance of the stub.

Fig. 2—Principle of double resonance.



B_c: Suseptance of the cavity.
B_x: Susceptance of the diode circuit.

Fig. 3—Principle of double resonance.

1.3. Experimental Results and the Consideration of Them

1) The case using stub resonance.

The circuit has been designed for a signal frequency of about 6 Gc and a pump frequency of 12 Gc. The separation between the two resonant frequencies corresponding to the signal and the idler can be varied by adjusting the depth and the position of the stub. Experiment has proved that the Q of the circuit is high and the bandwith is rather narrow. A gain of 32 dB for a bandwith of 2 Mc and a gain of 26 dB for a bandwith of 7 Mc are obtained when the pump power is 10 mW.

2) The case using the resonance of the diode circuit.

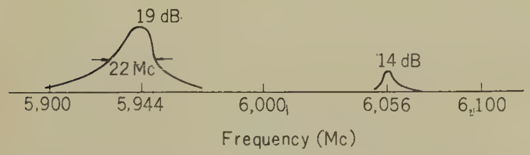


Fig. 4—Amplification performance by the doubly-resonant parametric amplifier.

In this case a comparatively broad-band characteristic have been obtained. An example is shown in Fig. 4. But in order to get a wide bandwidth the admittances of the circuits for the signal and the idler must be shifted slightly out of tune, and become inductive. When the both circuits are precisely tuned, amplification occurs with a very low level of pump power, but bandwidth is narrow and the amplifier is apt to oscillate.

Here, let us consider why a broad-band characteristic is obtained when the admittances of the circuits for the signal and the idler are shifted slightly out of tune. The equivalent circuit of a parametric resonator is shown as Fig. 5.(1) This figure shows that the resonant circuit for the w_1 is combined with the circuit for the idler w_2 through the nonlinear reactance of a semiconductor diode. Here the current and the voltage of the nonlinear reactance satisfy the following equation:(2)

$$\begin{pmatrix} I_1 \\ I_2{}^* \end{pmatrix} = \begin{pmatrix} jB_{11} & jB_{12} \\ jB_{21} & jB_{22} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2{}^* \end{pmatrix}$$
$$= \begin{pmatrix} 0 & j\omega_1 c_1 e^{j\theta_p} \\ -j\omega_2 c_1 e^{-j\theta_p} & 0 \end{pmatrix} \begin{pmatrix} V_1 \\ V_2{}^* \end{pmatrix} \tag{1}$$

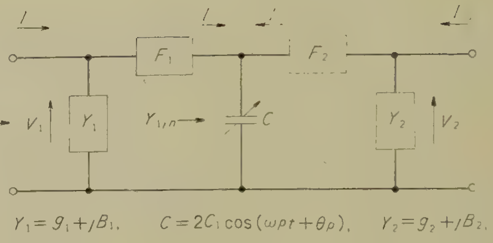
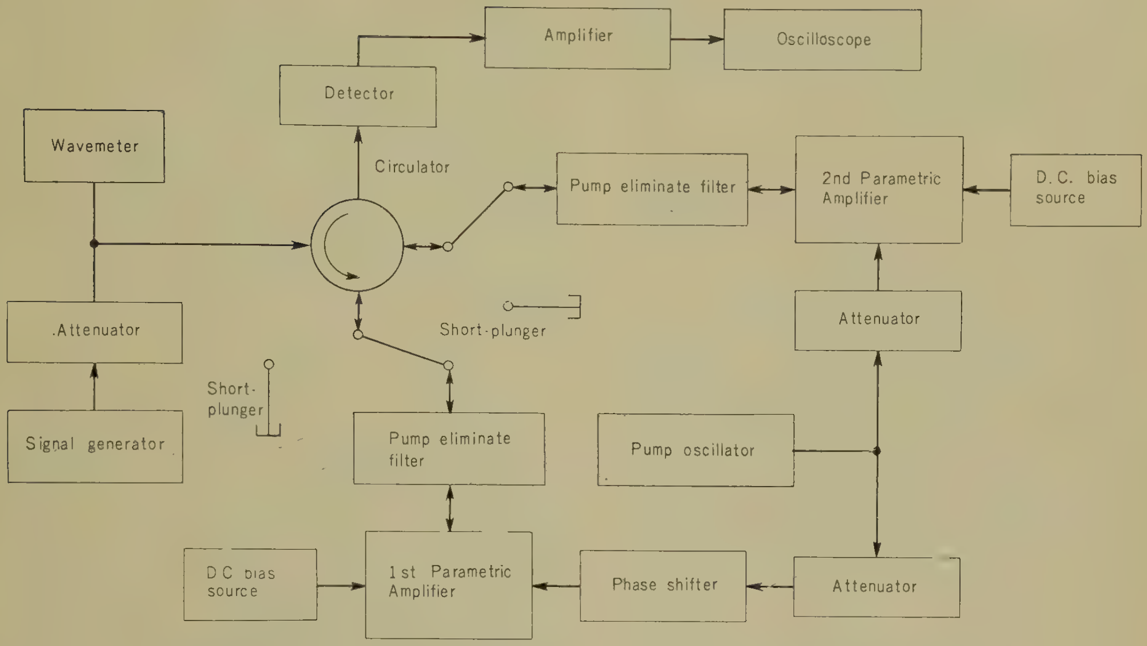


Fig. 5—Equivalent circuit of a parametric amplifier. F_1 and F_2 are ideal filters which let only I_1 and I_2 respectively pass.

When there is no idler input, the circuit admittance for ω_1 is:

$$Y_{1s} = g_1 + \frac{g_2 B_{12} B_{21}}{|Y_2|^2} + j\left(B_1 + \frac{B_2 B_{12} B_{21}}{|Y_2|^2} \right) \tag{2}$$

In this equation the second term of the real part expresses a negative conductance, and the second term of the imaginary part shows the effect of the imaginary part of the idler circuit connected through the nonlinear reactance. The experimental results are explained as follows: as B_{12} and B_{21} have opposite signs, the susceptances of the signal and the idler circuits, B_1 and B_2 in Equation (2), are compensated for by one another when they have the same sign. And the change of susceptance caused by the frequency deviation is not so steep as when each circuit is in tune. Therefore a relatively broad-band characteristic is obtained. Of course, the efficiency of amplification is reduced by the decrease of negative conductance, but it is



Eig. 6—The staggered amplifier arrangement connected in cascade.

compensated for by the increase of pump power.

2. Staggered Amplification

2.1. Circuit Arrangement and Experimental Results

The arrangement of an experimental circuit is shown in Fig. 6. Two resonant circuit parametric amplifiers are connected in cascade through a circulator. The staggered amplification of these circuits is not so simple as that of an ordinary cascade amplifier, because both the idler generated at the first stage and the output signal are fed to the second stage. Besides the shift of resonant frequencies of the two amplifiers as in the usual staggered operation, the variation of phase difference between the pumps of the two amplifiers can alter the total frequency characteristic. A relatively flat characteristic has been obtained by adjusting the pump phase. Fig. 7 shows one set of performance curves. In this case, the double resonance in each amplifier is obtained by making use of the resonance

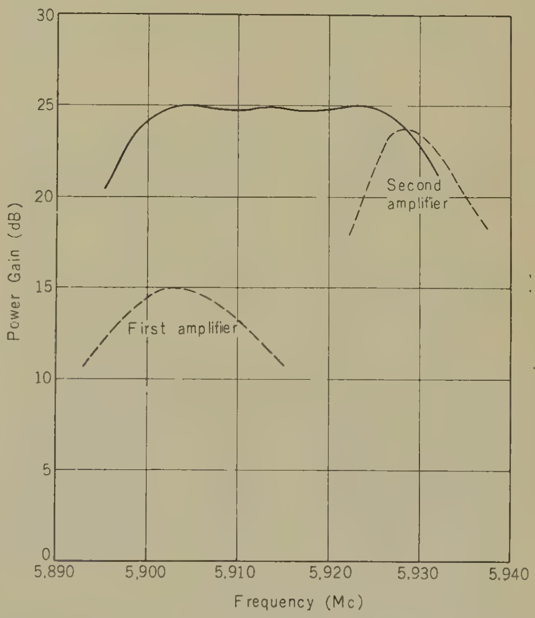
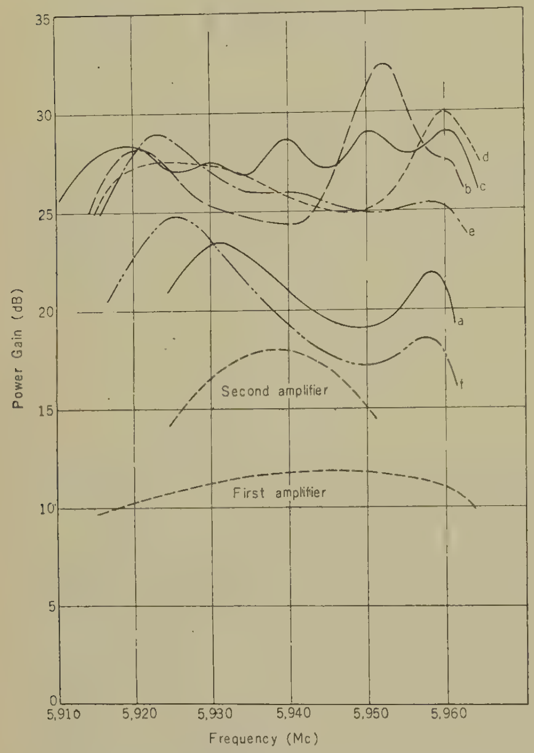


Fig. 7—Staggered amplification by two amplifiers connected in cascade.

of diode circuit itself. The gain is 25 dB for a bandwidth of 20 Mc over which the variation of the gain is less than 0.5 dB. The total pump power is 8 mW.



a: the case that the indication of the phase shifter θ_p is zero, b: $0.23\lambda_g$, c: $0.35\lambda_g$, d: $0.42\lambda_g$, e: $0.53\lambda_g$, f: $0.67\lambda_g$.

Fig. 8—Various frequency characteristics when θ_p is changed as a parameter.

Fig. 8 shows the effect of the pump phase on the frequency characteristic of gain. The pump input to each amplifier is 7 mW. For optimum operation the gain is 29 dB and the bandwidth 55 Mc.

2.2. Discussion

In this experiment the phase relation between the input signal, the input idler and the pump of the second amplifier plays a important part in determining the frequency characteristics. Therefore, first, consideration must be paid to the relation of the phase and the amplitude of the output signal and the idler of the first amplifier.

In Fig. 5, provided that currents I_{g_1} and I_{g_2}, which flow through g_1 and g_2, are the output currents of the signal and the idler respectively, and provided that the input idler is zero and g_1 is equal to g_2; the relation between I_{g_1} and I_{g_2} becomes:

$$\frac{I_{g_1}}{I_{g_2}{}^*}=-j\frac{Y_2{}^*}{\omega_2 c_1 e^{j\theta_p}} \qquad (3)$$

In the following discussion symbols for the second amplifier are primed to distinguish them from symbols relating to the first amplifier. For example, I_{1s} respresents the input signal current of the first amplifier, while I_{1s}' represents the input signal current of the second amplifier.

If the phase and amplitude relations between the output signal and idler current of the first amplifier are kept through the intermediate transmission line and the circulator, the relation between the input currents, I_{1s}' and I_{2s}', of the second amplifier are also expressed by Equation (3) by substituting I_{1s}' and I_{2s}' respectively for I_{g_1} and I_{g_2}.

In the second amplifier the action becomes somewhat complicated because of the input idler. The susceptance of variable capacitance Y'_{1in} for signal becomes:

$$Y'_{1in}=\frac{-\omega_1\omega_2 c_1'^2+j\omega_1 c_1' e^{j\theta_p'} Y_1'\left(\frac{I_{2s}'^*}{I_{1s}'}\right)}{Y_2'^*-j\omega_1 c_1' e^{jp\theta'}\left(\frac{I_{2s}'^*}{I_{1s}'}\right)} \qquad (4)$$

Now, if the voltage gain is defined as the ratio of the output voltage with pump and that without pump, it becomes:

$$G=\frac{1-jAe^{j(\theta_p'-\theta_1'-\theta_2'+\psi_2')}}{1-Be^{j(\psi_2'-\psi_1')}} \qquad (5)$$

where

$$A=\frac{\omega_2 c_1'}{Y_2'}\Big|\frac{I_{2s}'}{I_{1s}'}\cdot B=\frac{\omega_1\omega_2 c_1'^2}{Y_1'\ Y_2'},$$

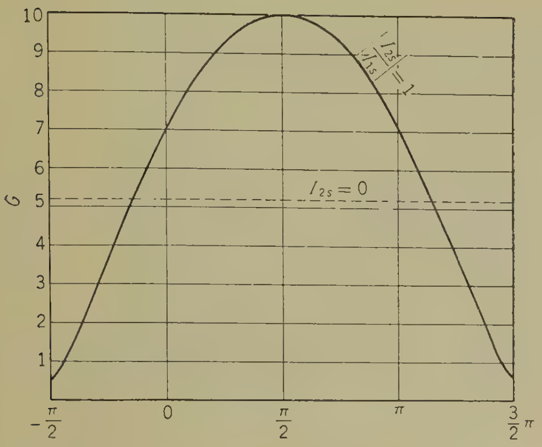


Fig. 9—Calculated values of the θ-voltage gain characteristic.

and θ_1', θ_2', ϕ_1', ϕ_2', are respectively the phase angle of I_{1s}', I_{2s}', Y_1', Y_2'. If θ is defined by the equation: $\theta=\theta_p'-(\theta_1'+\theta_2')$ the relation between θ and the gain G is, for example, as shown in Fig. 9. The curves are calculated under the assumption that Y_1' and Y_2' resonate to ω_1 and ω_2 respectively. The solid curve shows the case of $|I_{2s}'/I_{1s}'|=1$ and the dotted curve shows the case of $I_{2s}'=0$. The maximum difference between these two curves is nearly 6 dB. Thus the phase relation is a very large factor in determining the amplification characteristics of the second amplifier, while the output signal and the idler of the first amplifier, i.e., the input signal and idler of the second amplifier have the frequency characteristic of mutual relation as shown in Equation (3). Therefore the phase relation and its effect on the total frequency characteristic of amplification is not uniform over the amplification band. In the staggered operation, by adjusting θ_p so that the optimum point of θ may be brought to a proper frequency, the valley between two peaks of gain characteristic by the two amplifiers can be raised and a flat characteristic can be obtained. Fig. 8 illustrates this process.

Conclusion

This experiment has aimed at improving the bandwidth characteristics of resonant-circuit parametric amplifiers by staggered operation. It was proved that a carefully determined phase relation between the input signal, the input idler, and the pump of the second amplifier helps to flatten the gain-frequency characteristic in staggered operation. In other words, for a given allowable amplitude variation of the gain-frequency characteristic, a wider bandwidth can be obtained than in the ordinary staggered amplifier.

The noise figure of this staggered amplification is considered to be a little worse than that of a single stage amplifier owing to the addition of the noise of the second amplifier. But this added noise is not so large, so long as the gain of the first amplifier is high.

Acknowledgement

The author heartily thanks Dr. M. Kuwata and Dr. T. Hashimoto for their excellent advice.

References

(1) H. E. Rowe: Some General Properties of Nonlinear Elements, part II, Small Signal Theory, *Proc. I.R.E.*, **46**, No. 5, p. 850, 1958-5.
(2) Shigobumi Saito: Parametric Amplifier Circuits, *The Journ. of the Inst. of Electr. Comm. Engn. of Japan*, **42**, No. 6, p. 632, 1959-6.

* * * *

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 1-2, JANUARY-FEBRUARY, 1961

U.D.C. 621. 382. 2

Construction and Characteristics of Silver-Bonded Diodes*

Shoichi KITA†

A silver-bonded diode was invented and developed for use in parametric amplifiers. The diode was produced by the following process: The tip of a silver whisker containing a small amount of gallium was brought into contact with an n-type germanium wafer, and electrically formed. The impedance of the diode was measured with a coaxial type standing wave detector at room temperature and at 100° K. The minimum barrier capacitance was about 0.1 pF, the series resistance 10 ohms, and the cutoff frequency at breakdown voltage 200 GC.

1. Introduction

Several years ago the author noticed that negative resistance could be produced at microwave frequencies by applying pumping power, whose frequency was higher than the signal frequency, to a semiconductor diode. Using this negative resistance he succeeded in obtaining amplification and oscillation at microwave frequencies. This type of amplifier, now known as the parametric amplifier, has recently attracted much attention because of its low noise characteristics.

Large variation of barrier capacitance and high-cutoff frequency are required of parametric amplifier diodes. The author invented a silver-bonded diode which is constructed by bringing the tip of a silver-gallium whisker into contact with an n-type germanium wafer, and then electrically forming the resulting junction.

2. The Structure and the Method of Constructing Silver-Bonded Diodes

2.1. The Structure of the Silver-Bonded Diode

The silver-bonded diode is constructed by bringing the tip of a silver whisker containing gallium into contact with an n-type germanium wafer of about 0.02∼0.2 Ω-cm resistivity, and then by electrically forming the junction.

The actual diode is used in a holder, as shown in Fig. 1. Fig. 1 (a) shows a glass holder used mainly at 6 GC; Fig. 1 (b) shows a coaxial type holder, which because the length of the whisker can be made extremely short, is suitable for frequencies higher than 10 GC; while Fig. 1 (c) shows a waveguide-type holder in which the diode is mounted in a 10 GC-band waveguide which has been cut from the center of a metal plate.

The diode whisker is a piece of 50∼100μ silver wire containing 5∼6 % gallium. The tip of the whisker is shaped electrolytically in silver electrolytic solution until the diameter of the point decreases to several μ. The germanium wafer is etched electrolycally in chromium sulfate solution.

The tip of the silver whisker is brought into contact with the germanium wafer, and the junction is electrically formed by passing the discharge current of a capacitor through

---

* MS in Japanese received by the Electrical Communication Laboratory, April 18, 1960. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T., Vol. 9. No. 6, pp. 639-648, 1960.

† Radio Section.

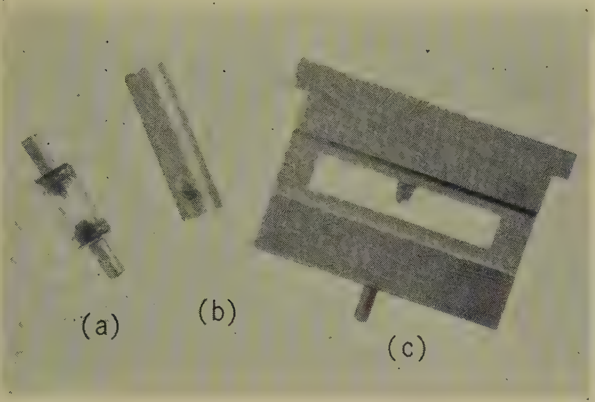


Fig. 1—Structure of silver-bonded diodes.

it.

The author devised an electrical forming process in which, instead of applying a high voltage in one step, the voltage is gradually increased from a very low value (IV) in a series of repeated steps.

2.2. The Method of Constructing Silver-Bonded Diodes

Owing to this electrical forming, the voltage-current characteristics of the diode change as indicated in Fig. 2. The figure in the circle indicates the number of the repetitions, and the number 0 indicates an unformed point contact. As the number of the electrical forming steps increases, the foreward current increases, but the reverse current does not show a large change, and the characteristics are improved.

3. Characteristics of the Silver-Bonded Diode

As the silver-bonded diode is mainly used as a component in 6 GC-Band parametric amplifiers, its characteristics were measured chiefly at 6 GC.

The impedance at 6 GC was measured with a coaxial-type standing-wave detector. The series resistance of the diode, the absolute value of the barrier capacitance, and the variation of barrier capacitance with applied voltage were calculated from this data.

The value of the barrier capacitance was measured at both 300°K and 100°K. The apparatus shown in Fig. 3 was used for the measurement at 100°K. The coaxial type standing-wave detector was set up vertically, and the diode mount was immersed and cooled in the liquid nitrogen contained in the vessel. The measured change of the barrier capacitance due to the bias voltage change is illustrated in Fig. 4. At the lower tempera-

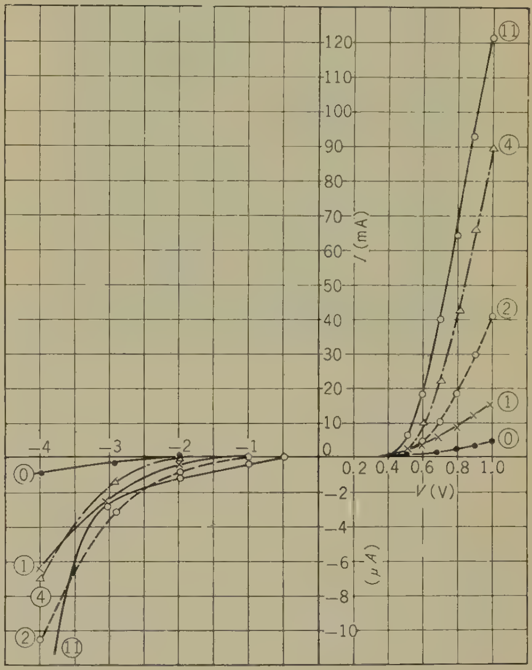


Fig. 2—Change of voltage-current characteristics of silver-bonded diodes due to repeated electrical forming.

ture, the absolute value of the barrier capacitance was lower than that at room temperature. The series resistance was also measured, but its value did not change from that at room temperature; therefore the cutoff frequency increased.

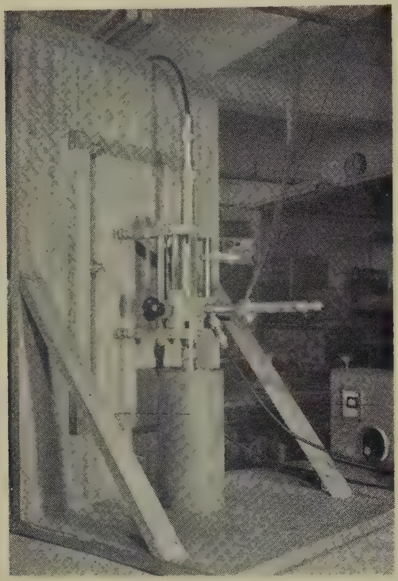


Fig. 3—Equipment for measuring the impedance of diodes at low temperatures.

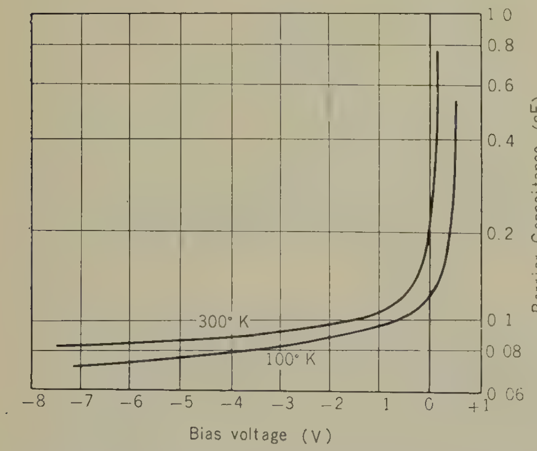


Fig. 4—Variation of barrier capacitance with bias voltage.

The silver-bonded diode developed for parametric amplifier use has a low foreward resistance and barrier capacitance lower than 0.3 PF. Therefore it is also considered to be suitable for use in computer circuits. When the diode is actually used in computer circuits, special attention should be paid to the transient characteristics which appear under pulse-voltage conditions. Fig. 5 shows the result of the measurement of the decay time with pulse-voltage applied to the silver-bonded diode, and to commercial SD 13, SD 16, SD 17 (a gold-bonded diode), and MA 301 (a silicon variable capacitance diode for TVAFC). The silver-bonded diode showed the shortest decay time, less than 10 mμs. No effects of hole storage were observed in the silver-bonded diode.

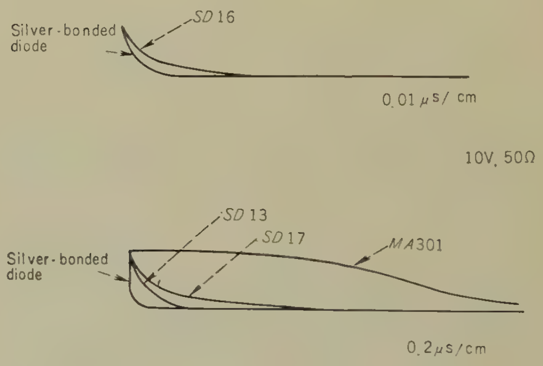


Fig. 5—Transient characteristics of diodes.

The diode was tested for resistance to mechanical shock by dropping it from a height of 1 meter on to a concrete floor, but no significant change in characteristics was observed. Also the effects of temperature and humidity were investigated by measuring the voltage-current characteristics both before and after immersion of the diode in boiling water, but again no significant changes were noted, and the high stability of the characteristics of the silver-bonded diode was confirmed.

Table 1

CONSTANTS OF DIODES

	R_s (Ω)	C_{min} (pF)	C_{0V} (pF)	$C_{min}{\cdot}R_s$ (pF·Ω)	f_c (Gc)
Silver-bonded* diode	5～10	0. 1～0. 2	0. 1～0. 3	0. 7～1. 2	130～230
1N23B*	36	0. 25	0. 35	9	18
1N23C	30	0. 15	0. 3	4. 5	36
Gold-bonded* diope	12	0. 12	0. 25	1. 4	112
Ge-Pt point-contact* diode	19	0. 11	0. 25	2. 1	79
Gold-bonded** diode	5	0. 6	2. 0	3. 0	53
Si-diffused** diode	4	0. 6	1. 6	2. 4	66

* Measured at 6 Gc.
** Parametric diode made in USA and measured at 900 Mc.

4. Comparison of the Characteristics of the Silver-Bonded Diode with Those of Other Diodes.

The characteristics of silver-bonded diodes, gold-bonded diodes, platinum and tungsten point contact germanium diodes, commercial 1 N 23 B's and 1 N 23 C's, and American parametric-amplifier diodes were compared.

Fig. 6 illustrates the voltage-current characteristics of these diode. The silver-bonded diodes have a larger foreward current than any other point contact diode and extremely small reverse current, proving the excellence of its characteristics.

In Table 1 the constants of the above diodes measured at 6 GC are tabulated; except that the American diodes were measured at 900 MC. In the table, R_s is the series resistance, C_{0V} is the barrier capacitance at 0 bias, C_{min} the barrier capacitance at the breakdown voltage, and f_c the cutoff frequency.

The table shows that there is no big difference in the values of C_{0V} and C_{min} between different point contact diodes, (i.e. all

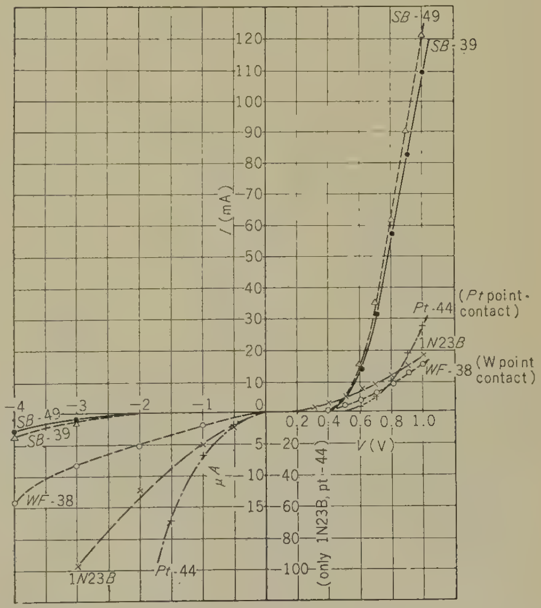


Fig. 6—Voltage-current characteristics various of diodes.

values are in the range between 0.1～0.3 pF.) The R_s of the silver-bonded diode is 5～10 Ω, whereas those of the other point contact diodes are as high as 10～40 Ω. Therefore the values of the cutoff frequency of these diodes, excepting the gold-bonded diode, are from one-third to one-tenth those of the silver-bonded diode. For the American diodes, the values of R_s of both the bonded type and the diffused type diodes are smaller than 5 Ω, but the values of C_{0V} and C_{min} are 0.6～2 pF, which are about ten times those of the silver-bonded diode. Therefore the value of the cutoff frequency of the American diodes is 50～60 GC, which is about one-half to one-third that of the silver-bonded diode.

The reason for the extremely high cutoff frequency of the silver-bonded diode is considered to be as follows: The smaller the contact area of the diode becomes, the higher the cutoff frequency. Yet in the diffused type, the contact area cannot be made very small, whereas in the point-contact type, the effect of the stray capacitance prevents the cutoff frequency from becoming very high if the contact area becomes too small. The result is that there is an appropriate contact area to obtain the maximum cutoff frequency in point-contact diode. In the silver-bonded diode, thanks to the repeated forming process, it is possible to control the contact area at will. Therefore by obtaining an appropriate contact area, the silver-bonded diode can obtain the highest cutoff frequency.

5. Conclusion

The method of constructing the silver-bonded diode invented by the author has been carefully examined so as to be able to produce diodes with the best possible characteristics for parametric amplifier use.

The diode series resistance is less than 10 Ω, barrier capacitance is less than 0.3 pF, and cutoff frequency is more than 200 GC. Its characteristics do not deteriorate at temperatures as low as 100°K. This fact proves that the diode is suitable for a parametric amplifier which should be cooled to the temperature of liquid nitrogen in order to obtain a low noise figure.

Using this diode, 4 GC, 6 GC and 11 GC band parametric amplifiers were experimentally constructed. The minimum noise figure for single sideband reception was about 5 dB in 4 GC(3) and 6 GC bands, and about 6.5 dB in the 11 GC band.(4) When used as a transmitting frequency converter(5) and as a hamonic generator, they proved to be more than several dB more efficient than other diodes.

Silver-bonded diodes are still being studied. It is expected that a higher frequency usage than that described above will be attained.

References

(1) Kita and Fujii: "Microwave Amplification by Use of Crystal Diode" in *ECL International Rept.*, Feb., 1954. *Annual Convention*, Inst. of Elec. Commun. Engrs., Japan, Oct., 1954.

(2) Kita, Sugiyama and Okajima. "Properties of Silver-Bonded Diode for a Parametric Amplifier," 1959 Electron Devices Meeting, Oct., 1959.

(4) Kita and Obata: "An X-band Parametri Amplifier Using a Silver-bonded Diode," *Proc. I.R.E.*, G **48**, 9, pp. 1651–1652, 1960.

(5) Kita, Sanpei and Okajima: "Transmitting frequency Converter in Which Gold or Silver-Bonded Diode is Used," *Rep. ECL.*, G **6**, 11, pp. 415–420, 1958.

(6) Kita and Sugiyama: "The Nonlinear Barrier Capacitance of Silver-Bonded Diodes," *Rev. ECL.*, G 8, 3–4, pp. 112–118, 1960.

* * * *

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 1-2, JANUARY-FEBRUARY, 1961

U.D.C. 678.012:621.928.5:542.65

Determination of Molecular Distribution of High Polymers by Fractionation*

Hiroshi OKAMOTO†

Several questionable assumptions are unavoidable in the determination of molecular weight distribution of high polymers by fractionation, and the reliability of the measured distribution is problematical.

The author considers the theoretical bases of two fractionation methods, i.e., Spencer's method and the successive precipitation method (SPM). Solving a set of simultaneous integral equations derived from Flory's theory, the author proposes a modification of Spencer's method. This result is found to be in accord with the numerical calculations. A criterion which assures us of the correctness of the distribution measured by SPM is also given in this paper. (Section 2).

Introduction

Molecular weights of high polymers may be established by appropriate physical measurements on dilute solutions, such as light scattering or the osmotic method. These methods are all absolute, because, if the experiments are carried out under ideal conditions, no error is introduced into the results.

In the case of determination of molecular weight distribution of high polymers by fractionation, however, such a situation does not hold. Because the procedures of fractionation are based on many questionable assumptions, there must be unavoidable errors in the resultant distribution curves.

Billmeyer and Stockmeyer,(1) who had studied Spencer's fractionation method,(2) came to a rather negative conclusion about this method from many experimental results. Other studies(3)(4)(5)(6) of the above mentioned problem have been made only through numerical calculations, but it seems difficult to obtain any general conclusions from them.

Because of these circumstances, the theoretical bases of the two fractionation methods, Soencer's method and successive fractionation method are reconsidered. A modification of Spencer's method and fractionation conditions which minimize errors are presented.

1. Spencer's Method

1.1. General Considerations

The procedure of Spencer's method is as follows: A series of polymer solutions is prepared. Each solution is separated into two phases by successive variation of the nonsolvent content or of the temperature of the solution. The quantities actually measured are the weight fraction of the precipitated polymer $W(s)$ and its weight-average molecular weight $m(s)$. To simplify the problem, Spencer proposed the artificial assumption that all the polymer above a certain molecular weight $x(s)$ is in the precipitate and that all below $x(s)$ is in the solution phase. Here, "s" is a parameter which represents the solvent power of the solution. It has one to one correspondence to the non-solvent content

---

* MS in English received by the Electrical Communication Laboratory on 15 September, 1960. Originally published in Journ. of the Phys. Soc. of Japan, Vol. 14, No. 10, Oct., pp. 1388–1396, 1960.

† High-Polymers Research Section.

or to the temperature of the solution.

According to the above assumption, if $f(x)$ is the distribution function of the molecular weight,

$$W(s)=\int_{x(s)}^{\infty} f(x)dx \qquad (1.1.1)$$

$M(s)$, the product of $m(s)$ and $W(s)$, may be written as

$$M(s)=m(s)W(s)=\int_{x(s)}^{\infty} xf(x)dx \qquad (1.1.2)$$

From Eqs. (1.1.1) and (1.1.2), the following equation is obtained,

$$\frac{dM}{dW}=x(s) \qquad (1.1.3)$$

In the next stage of the procedure, $m(s)\times W(s)$ versus $W(s)$ is plotted, and a smooth curve is drawn through the points. The curve is differentiated to give the corresponding values of $x(s)$ and $W(s)$. A plot of $W(s)-1$ versus $x(s)$ is the integral distribution curve.

It is an obvious fact, however, that neither Eq. (1.1.1) nor (1.1.2) is exact. There must be a continuous function $K(s,x)$ which determines the phase separation of the solution. The product $K(s,x)f(x)$ is proportional to the distribution function of the precipitated polymer. $W(s)$ and $M(s)$ are given by

$$W(s)=\int_{0}^{\infty} K(s,x)f(x)\,dx \qquad (1.1.4)$$

$$M(s)=\int_{0}^{\infty} K(s,x)xf(x)\,dx \qquad (1.1.5)$$

For a single solvent solution, the theory of Flory(7) or Scott(8) shows that $K(s,x)$ is given by

$$K(s,x)=\frac{1}{1+R(s)\exp[-\sigma(s)x]} \qquad (1.1.6)$$

where $R(s)$ is the ratio of the volume of the solution phase to that of the precipitated phase, and $\sigma(s)$ is a function of the interaction parameter and the composition of the solution.

For a single solvent solution, Flory's theory gives $\sigma(s)$ as

$$\sigma=v\left(1-\frac{1}{r_n}\right)-v'\left(1-\frac{1}{r_n'}\right)+\chi\{(1-v)^2-(1-v')^2\}. \qquad (1.1.7)$$

In this equation, v and v' are the volume concentration of the polymer in the solution and the precipitated phase respectively; r_n and r_n' are the number-average molecular weights of the polymer in respective phases; and χ is the interaction parameter between the solvent molecule and the polymer segment.

Even for a mixed-solvent solution (good solvent and nonsolvent), as long as the same order of approximation as Flory's is used, an equation identical with Eq. (1.1.6) can be derived. In this case, the expression for $\sigma(s)$ is more complicated than Eq. (1.1.7).

But it is needless in the author's case to know the detailed form of $\sigma(s)$ such as given by Eq. (1.1.7). Because the quantities obtainable directly from the experiments are $W(s)$ and $M(s)$, it is proposed that Eqs. (1.1.4) and (1.1.5) be regarded as a set of simultaneous integral equations with respect to two unknown functions $\sigma(s)$ and $f(x)$. $W(s)$, $M(s)$ and $R(s)$ are regarded as three known functions the variable of which is "s". This proposal requires the measurement of $R(s)$, which is trouble-some but unavoidable.

The approximate solution of this integral equation will be described in the next paragraph.

1.2. Solusion of the Proposed Simultaneous Integral Equations

We write down again the fundamental equations as

$$W(s)=\int_{0}^{\infty}\frac{1}{1+R(s)\exp[-\sigma(s)x]}f(x)dx \qquad (1.2.1)$$

$$M=m(s)W(s)$$

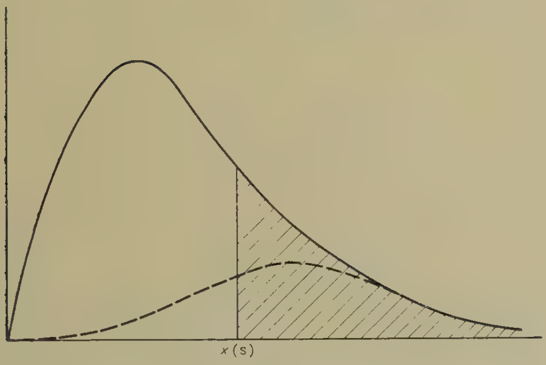
$$=\int_0^\infty \frac{1}{1+R(s)\exp[-\sigma(s)x]}x^\alpha f(x)dx. \tag{1.2.2}$$

When $\alpha=1$, the last equation is identical with Eq. (1.1.5), and $m(s)$ is the weight-average molecular weight. If the distribution curve is plotted against intrinsic viscosity, α is identical with that encountered in the intrinsic viscosity-molecular weight relationship,

$$[\eta]=kx^\alpha. \tag{1.2.3}$$

In this case, $m(s)$ is given by

$$m(s)=\frac{1}{k}[\eta]. \tag{1.2.4}$$



——the M.W. distribution of the original polymer.
……the M.W. distribution of the precipitated polymer. The area below the dotted line is equal to the shaded area.

Fig. 1—Schematic illustration of the phase separation.

There must be a value $x(s)$, and the integral of $f(x)$ from $x(s)$ to ∞ is equal to $W(s)$.

$$W(s)=\int_{x(s)}^\infty f(x)dx. \tag{1.2.5}$$

The $x(s)$ in this paragraph must not be confused with the one which appears in Spencer's approximation, Eq. (1.1.1). The $x(s)$ in this paragraph has only a mathematical significance, and never has a physical meaning as the one in Eq. (1.1.1). This is illustrated in Fig. 1, where the area below the dotted line is equal to the shaded area. $\Delta(s)$ is defined as

$$\Delta(s)=M(s)-\int_{x(s)}^\infty x^\alpha f(x)dx. \tag{1.2.6}$$

Then we can find that dM/dW is given by

$$\frac{dM(s)}{dW(s)}=x(s)^\alpha-\frac{1}{f\{x(s)\}}\cdot\frac{d\Delta(s)}{dx(s)}. \tag{1.2.7}$$

The second term of the right-hand side of the above equation gives the error of the unmodified Spencer's method.

$K(s,x)$ has an inflection point at $x=y(s)$, where

$$y(s)=\frac{\log R(s)}{\sigma(s)}. \tag{1.2.8}$$

The value of $K(s,x)$ at the inflection point is 1/2, and the following relation is always satisfied.

$$\int_{y(s)}^{x(s)} f(x)dx+\int_0^{y(s)}K(s,x)f(x)dx$$
$$-\int_{y(s)}^\infty\{1-K(s,x)\}f(x)dx=0. \tag{1.2.9}$$

The functions $K(s,x)$ and $f(x)$are expanded as

$$K(s,x)=\sum_{\nu\geqslant0}(-)^\nu\left\{\frac{1}{R(s)}\right\}^{\nu+1}\exp[(\nu+1)\sigma(s)x]$$
$$\text{for}\quad x<y(s) \tag{1.2.10}$$

$$1-K(s,x)=\sum_{\nu\geqslant0}(-)^\nu R(s)^{\nu+1}\exp[-(\nu+1)\sigma(s)x]$$
$$\text{for}\quad x>y(s) \tag{1.2.11}$$

$$f(x)=\sum_n\frac{1}{n!}\left(\frac{d^nf}{dx^n}\right)_s\{x-y(s)\}^n \tag{1.2.12}$$

where $(d^nf/dx^n)_s$ is the value of d^nf/dx^n at $x=y(s)$. Introducing these equations into Eq.

(1.2.9) and using term-by-term integration we arrive at

$$
\begin{aligned}
&\{x(s)-y(s)\}f(\xi_s)-2\sum_{\nu\geqslant 0}C_{2\nu+2}\left(\frac{d^{2\nu+1}f}{dx^{2\nu+1}}\right)_s\\
&\times\left(\frac{y(s)}{\log R\,(s)}\right)^{2\nu+2}-\sum_{\mu\geqslant 0}(-)^{\mu}D_{\mu+1,s}\left(\frac{d^{\mu}f}{dx^{\mu}}\right)_s=0
\end{aligned}
\tag{1. 2. 13}
$$

where,

$$C_n=\sum_{\nu\geqslant 0}\frac{(-)^{\nu}}{(\nu+1)^n},$$

$$
\begin{aligned}
D_{n,s}=&\frac{(-)^{n-1}}{(n-1)!}\frac{\partial^{n-1}}{\partial\sigma^{n-1}}\Biggl\{\frac{1}{\sigma}\\
&\times\sum_{\nu\geqslant 0}\frac{(-)^{\nu}}{(\nu+1)^n}\exp\left[-(\nu+1)\sigma y\right]\Biggr\},
\end{aligned}
$$

$$\int_{y(s)}^{x(s)}f(x)dx=\{x(s)-y(s)\}f(\xi_s)$$

$$y(s)\gtreqless\xi_s\gtreqless x(s).$$

By differentiating Eq. (1.2.13) with respect to $x(s)$, the following equation is obtained:

$$
\begin{aligned}
&\Biggl[f\{y(s)\}+2\sum C_{2\nu+2}\Biggl\{\frac{2\nu+2}{y(s)}\left(\frac{d^{2\nu+1}f}{dx^{2\nu+1}}\right)_s\\
&+\left(\frac{d^{2\nu+2}f}{dx^{2\nu+2}}\right)_s\Biggr\}\left\{\frac{y(s)}{\log R(s)}\right\}^{2\nu+2}\Biggr]\\
&\times\frac{dy(s)}{dx(s)}=f\{x(s)\}+2\Biggl[\sum_{\nu\geqslant 0}\frac{2\nu+2}{y(s)}C_{2\nu+2}\left(\frac{d^{2\nu+1}f}{dx^{2\nu+1}}\right)_s\\
&\times\left\{\frac{y(s)}{\log R(s)}\right\}^{2\nu+3}\Biggr]\frac{d\log R(s)}{dx(s)}\\
&-\frac{d}{dx(s)}\left\{\sum(-)^{\mu}D_{\mu+1,s}\left(\frac{d^{\mu}f}{dx^{\mu}}\right)_s\right\}.
\end{aligned}
\tag{1. 2. 14}
$$

From Eq. (1.2.6) and a procedure similar to that of deriving Eq. (1.2.13), the following equations are obtained:

$$
\begin{aligned}
\Delta(s)=&\int_{y(s)}^{x(s)}g(x)dx-2\sum C_{2\nu+2}\left(\frac{d^{2\nu+1}g}{dx^{2\nu+1}}\right)_s\\
&\times\left\{\frac{y(s)}{\log R(s)}\right\}^{2\nu+2}-\sum_{\nu\geqslant 0}(-)^{\mu}D_{\mu+1,s}\left(\frac{d^{\mu}g}{dx^{\mu}}\right)_s
\end{aligned}
\tag{1. 2. 15}
$$

$$g(x)=\{x^{\alpha}-y^{\alpha}(s)\}f(x).$$

To derive the second term of the right-hand side of Eq. (1.2.7), we must differentiate Eq. (1.2.15) with respect to $x(s)$.

$$
\begin{aligned}
\frac{d\Delta(s)}{dx(s)}=&\{x^{\alpha}(s)-y^{\alpha}(s)\}f\{x(s)\}\\
&-2\sum_{\nu\geqslant 0}C_{2\nu+2}\Biggl\{\frac{2\nu+2}{y(s)}\left(\frac{d^{2\nu+1}g}{dx^{2\nu+1}}\right)_s\\
&+\frac{d}{dy}\left(\frac{d^{2\nu+1}g}{dx^{2\nu+1}}\right)_s\Biggr\}\left\{\frac{y(s)}{\log R(s)}\right\}^{2\nu+2}\frac{dy(s)}{dx(s)}\\
&+2\sum_{\nu\geqslant 0}\frac{2\nu+2}{y(s)}C_{2\nu+2}\left(\frac{d^{2\nu+1}g}{dx^{2\nu+1}}\right)_s\\
&\times\left\{\frac{y(s)}{\log R(s)}\right\}^{2\nu+3}\frac{d\log R(s)}{dx(s)}\\
&-\frac{d}{dx(s)}\left[\sum_{\mu\geqslant 0}(-)^{\mu}D_{\mu+1,s}\left(\frac{d^{\mu}g}{dx^{\mu}}\right)_s\right].
\end{aligned}
\tag{1. 2. 16}
$$

Now, we must consider the convergence of the various series shown above.

If there exists a valve Φ_n which satisfies the relations below for any x in some range A,

$$\frac{\left|x^n\dfrac{d^nf}{dx^n}\right|}{f(x)}\leq\Phi_n, \tag{1. I}$$

$$x\varepsilon A$$

and if $\log R$ satisfies the relation

$$\frac{\Phi_{n+1}}{\Phi_n}<\log R, \tag{1. II}$$

then the higher terms of the series expansions Eqs. (1.2.13) and (1.2.15) may be neglected.

By eliminating $dy(s)/dx(s)$ from Eq. (1.2.16) by the use of Eq. (1.2.14), the following equation is obtained from Eqs. (1.2.7) and (1.2.16):

$$\frac{dM}{dW}=x^{\alpha}(s)+2C_2\alpha x^{\alpha-1}(s)\Bigg(\Bigg\{\frac{1}{f(x(s))}-\frac{1}{f(\xi_s)}\Bigg\}\Bigg(\frac{df}{dx}\Bigg)_s+\frac{\alpha+1}{y(s)}\Bigg]\Bigg\{\frac{y(s)}{\log R(s)}\Bigg\}^2$$

$$-4C_2\alpha y^{\alpha-2}(s)\frac{f(y(s))}{f(x(s))}\cdot\Bigg\{\frac{y(s)}{\log R(s)}\Bigg\}^3\times\frac{d\log R(s)}{dx(s)}+\cdots\cdots. \qquad (1.\,2.\,17)$$

It can also be found from Eq. (1.2.13) that the difference between $x(s)$ and $y(s)$ or $x(s)$ and ξ_s may be placed outside of our consideration as far as we content ourselves with an approximation of the same order as Eq. (1.2.17). Thus Eq. (1.2.17) can be simplified further as

$$\frac{dM}{dW}=x^{\alpha}(s)\Bigg(1+C_2\Bigg\{\frac{2\alpha(\alpha+1)}{(\log R(s))^2}-\frac{4\alpha^2}{(\log R(s))^3}\cdot\frac{d\log R(s)}{d\log x(s)}\Bigg\}\Bigg). \qquad (1.\,2.\,18)$$

It is important that the unknown function $f(x)$ does not appear in Eq. (1.2.18). As the value of $\log R$ becomes larger, the terms higher than the second become negligible. Finally, we arrive at

$$\frac{dM}{dW}=x^{\alpha}(s). \qquad (1.\,2.\,19)$$

That is, the zeroth order approximation of our treatment is in accord with the unmodified Spencer's method. In Eq. (1.2.18) the unknown differential $d\log x(s)$ is in the right-hand side, so Eq. (1.2.18) is a differential equation for $x(s)$. The approximate solution of this equation is obtained by replacing $x(s)$ with $\{d\log dM/dW\}/\alpha$. That is,

$$x(s)^{\alpha}=\frac{dM/dW}{1+C_2\left(\frac{2\alpha(\alpha+1)}{(\log R(s))^2}-\frac{4\alpha^2}{(\log R(s))^3}\frac{d\log R(s)}{d\log(dM/dW)}\right]} \qquad (1.\,2.\,20)$$

From Eq, (1.2.20), $x(s)$ can be determined in terms of the observable quantities R and dM/dW. C_2 is easily found to be $\pi^2/12$. The first order approximation of the integral distribution $I(x)$ is finally obtained by the help of Eq. (1.2.5)

$$I\{x(s)\}=1-W(s).$$

1.3. Discussion

Now a brief description of Matumoto's numerical calculations(6) will be made. First, he assumed $f(x)$ to be several kinds of known functional forms such as Schulz's distribution function. The values of $\sigma(s)$s were assumed to be a series of properly determined numerical values. The values of $R(s)$s were assumed to be constant throughout. The series of values of $W(s)$ and $M(s)$ were numerically calculated by the use of Eqs. (1. 2,1) and (1.2.2). Then, making use of these values of $W(s)$ and $M(s)$, he calculated the integral distribution following Spencer's method. Some of his calculations are shown in Table 1 and Fig. 2. It is shown in Fig. 2 that the integral distribution curves obtained by usual Spencer's method shift to the right side of the true one. Our solution Eq. (1.2.18) or Eq. (1.2.20) can explain this aspect. In Matumoto's calculations, the $R(s)$s were assumed to be constant throughout, so the second term in the square blackets is considered to be zero. Thus we can verify that

$$(dM)/(dW)>x.$$

This is a theoretical proof of the feature of the distribution curves shown in Fig. 2.

Matumoto also obtained the curves for the box type and the doubled-peak type distributions. For the box type distribution, his calculated curve seems to lie to the right of

Table 1

DATA OF THE NUMERICAL CALCULATIONS MADE BY MATUMOTO

$R=20$	$\sigma\times10^3$	2	3	5	10	20
	$W(s)$	0.786	0.654	0.411	0.153	0.268
	$M(s)$	520	388	190	48.9	5.88
	dM/dW	1050	842	580	335	175
$R=1000$	$\sigma\times10^3$	5	7	10	15	20
	$W(s)$	0.870	0.720	0.513	0.284	0.171
	$M(s)$	566	403	232	94.7	45.0
	dM/dW	1270	930	685	460	355

Table 2

CONDITIONS OF THE NUMERICAL CALCULATIONS : BOX TYPE DISTRIBUTION

R	1000	800	640	512	410	328	262	210	168	134
$\sigma\times10^3$	3.498	3.960	4.478	5.059	5.704	6.423	7.223	8.115	9.101	10.18

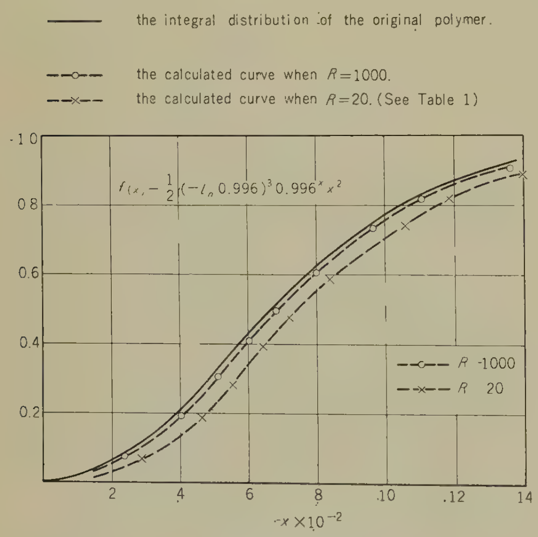


Fig. 2—Some of the results of Matumoto's numerical calculations.

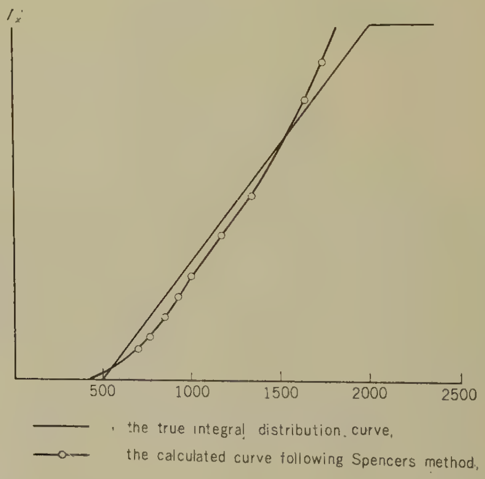


Fig. 3—Numerical calculation for box type distribution. (See Table 2.)

the true curves. We tried similar and more detailed numerical calculations for the box type distribution and found that the calculated curve increases rapidly and moves from the right side to the left side of the true curve near the end of the distribution on the higher molecular weight side (Table 2 and Fig. 3 R was not assumed to be constant).

It can be explained at least partly, why such an irregularity arises. We have found in Section 1.2 that conditions (1. I) and (1. II) must be satisfied to make Eq. (1.2.18) correct. The box type distribution function $f(x)$ does not satisfy condition (1. I), because it has two points of discontinuity at both ends of the distribution.

Though it may be too bold, the following remarks will also be added:

From the statistical analysis of their experiments, Billmeyer and Stockmeyer(1) concluded that Spencer's method cannot yield a very detailed description of the molecular weight distribution. They considered three types of distribution curves, a rectangular box, a triangle, and a parabola. According to our consideration, none of these distribution curves satisfy conditions (1. I) and (1. II), because they have two or three points of discontinuity. The application of Spencer's method is questionable in these cases, and this may be one of the important reasons for the failure of Spencer's method.

Here arises an important question, i.e., how to know whether or not $f(x)$ satisfies conditions (1. I) and (1. II). At the beginning of the experiment, nothing is known about $f(x)$, and R can not be determined properly to satisfy condition (1. II). One of the possible methods of answering this question is to check whether the zeroth or the first order approximations is satisfied by conditions (1. I) and (1. II). If they are satisfied, we can almost safely conclude (not always strictly) that the true distribution curve also satisfies the two conditions. If the approximated function does not satisfy the two conditions, it can strictly be said that the obtained distribution function never represents the true function, and we are obliged to recommence the fractionation with larger R.

A short remark must be added for the measurement of the average molecular weight. The use of Z or higher order average molecular weight must be avoided. Because for these averages the theory developed in Section 1.2 does not hold and the final result, Eq. (1.2.20), cannot be used.

2. Successive Fractionation

2.1. General Consideration

In the preceding section, the theoretical basis of Spencer's method has been considered and a more reliable method has been proposed. However, the method which is thought to be most widely employed now is the successive fractionation method, especially the successive precipitation method (S.P.M.). There are several questionable assumptions in the usual procedure of S.P.M. They are:

The polymer sample whose molecular weight distribution is to be determined is dissolved in an appropriate solvent or solvent mixture. By means of lowering the temperature or adding a precipitant, its solvent power is lowered slowly until two-phase separation occurs. Then the precipitated phase is taken out of the apparatus. The ratio of the weight of the precipitated polymer to that of the whole polymer W_1 and the average molecular weight of the precipitated polymer m_1 are measured, and similar procedures are repeated $n-1$ times. Finally W_n and m_n of the remaining solution are measured. Usually the integral molecular weight distribution curve is assumed to be

$$I(m_i)=\sum_{\lambda=0}^{\lambda=n-i-1} W_{n-\lambda}+\frac{1}{2}W_i. \qquad (2.1.1)$$

As already pointed out by Schulz,(3) the actual molecular weight distribution of each fraction overlaps each other considerably. Because this overlapping is apparently neglected in Eq. (2.1.1), the reliability of the obtained distribution is very doubtful. The influence of this neglect upon the results has been numerically calculated by G. V. Schulz (3) and Matumoto and Oyanagi,(5) but no general conclusions can be drawn from these numerical calculations.

The questions, i.e., how this negligence affects the result and how the experiments should be performed to minimize the errors, have not yet been solved.

Such questions will be answered in the following by similar treatments as was described in Section I.

2.2. A Criterion for Fractionation

In Fig. 4, a graphical scheme of a successive fractionation is shown.

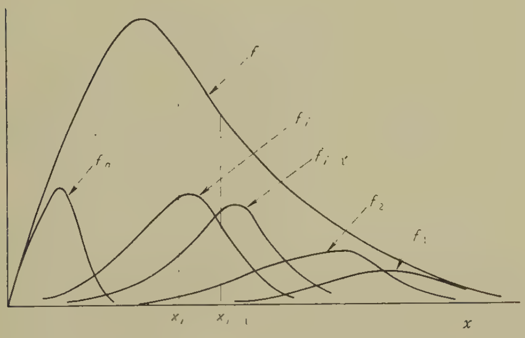


Fig. 4—Schematic representation of successive precipitation method. f: the distribution of the original polymer. $f_1, f_2 \cdots\cdots f_n$: the distribution of the fractionated polymer. The area below the curve f between x_i and x_{i-1} is equal to the area below the curve f_i.

The original polymer whose distribution $f(x)$ is to be determined is fractionated into the fractions (1), (2),$\cdots\cdots$(n). The distributions of (1), (2),$\cdots\cdots$(n) are shown by figures $f_1, f_2, \cdots\cdots f_n$. The function f_i is integrated as

$$\int_0^\infty f_i(x)\,dx = W_i. \tag{2.1.2}$$

"x_{i-1}" is the lower limit of the integral of $f(x)$. The value of this integral is given by

$$\int_{x_{i-1}}^\infty f(x)\,dx = \sum_{\lambda=1}^{i-1} W_\lambda. \tag{2.1.3}$$

Consequently, we have

$$\int_{x_i}^{x_{i-1}} f(x)\,dx = W_i. \qquad \text{(See Fig. 4.)} \tag{2.1.4}$$

M_i is defined as

$$M_i = m_i W_i \tag{2.1.5}$$

$$= \int_0^\infty x f_i(x)\,dx \tag{2.1.6}$$

Δ_i is defined as the difference

$$\Delta_i = M_i - \int_{x_i}^{x_{i-1}} x f(x)\,dx. \tag{2.1.7}$$

If
$$\Delta_i \ll \int_{x_i}^{x_{i-1}} x f(x)\,dx \tag{2. I}$$

then m_i is in the range

$$x_i < m_i < x_{i-1}.$$

The integration range $[x_i, x_{i-1}]$ will be called d_i hereafter. It is important for us to notice in this stage the following consideration:

If d_i is so small that $f(x)$ can be seen constant in this range, then

$$m_i = \frac{1}{2}(x_{i-1} + x_i)$$

$$\int_{m_i}^\infty f(x)\,dx = \sum_{\lambda=1}^{i-1} W_\lambda + \frac{1}{2} W_i.$$

The usual procedure of Eq. (2.1.1) is found to be adequate in this case. Therefore there are two important factors in the fractionation, one being to make d_i as small as possible and the other being to make Δ_i satisfy condition (2. I).

Generally speaking, the former condition requires us to make the amount of each fraction as small as possible or to make the number of fractions as large as possible. But the latter condition is not so obvious. In the following paragraph, it will be clarified.

2.3. Calculation of the Fractionation Condition

There are n-1 separation functions, which are similar to those used in the preceding section, for $n-1$ fractionation step.

$$K_i(x)=\frac{1}{1+R_i\exp[-\sigma_i x]}. \qquad (2.3.1)$$

$$i=1, 2, \cdots\cdots, n-1$$

The distribution function of the ith fraction is proportional to

$$K_i \prod_{\lambda=1}^{i-1}(1-K_\lambda)f(x).$$

W_i is written as

$$W_i=\int_0^\infty K_i(x)\prod_{\lambda=1}^{i-1}\{1-K_\lambda(x)\}f(x)dx. \qquad (2.3.2)$$

Considering that distribution is often plotted against intrinsic viscosity,

$$M_i=\int_0^\infty K_i(x)\prod_{\lambda=1}^{i-1}\{1-K_\lambda(x)\}x^a f(x)dx. \qquad (2.3.4)$$

Subsequently, the following relations can also be found:

$$1-\sum_{\lambda=1}^{i} W_\lambda=\int_0^\infty \prod_{\lambda=1}^{i}\{1-K_\lambda(x)\}f(x)dx \qquad (2.3.5)$$

$$M_0-\sum_{\lambda=1}^{i} M_\lambda=\int_0^\infty \prod_{\lambda=1}^{i}\{1-K_\lambda(x)\}x^a f(x)dx \qquad (2.3.6)$$

where $M_0=\int_0^\infty x^a f(x)dx.$

We define x_i as the upper limit of an integral of $f(x)$ whose value is given by the following equation:

$$\int_0^{x_i} f(x)\,dx=\int_0^\infty \prod_{\lambda=1}^{\infty}\{1-K_\lambda(x)\}f(x)dx. \qquad (2.3.7)$$

The following relations follow from the above equation:

$$1-\sum_{\lambda=1}^{i} W_\lambda=\int_0^{x_i} f(x)dx \qquad (2.3.8)$$

$$W_j=\int_{x_j}^{x_{j-1}} f(x)dx. \qquad (2.3.9)$$

The following equation stands for any y_i, but it is considered to be the inflection point of K_i.

$$\int_{x_i}^{y_i} f(x)dx=\int_0^{y_i}\left\{1-\prod_{\lambda=1}^{i-1}(1-K_\lambda)\right\}fdx+\int_0^{y_i}K_i\prod_{\lambda=1}^{i-1}(1-K_\lambda)f\,dx-\int_{y_i}^{\infty}(1-K_i)\prod_{\lambda=1}^{i-1}(1-K_\lambda)fdx. \qquad (2.3.10)$$

The following relations can be proved by procedures analgous to those which were used to derive Eqs. (1.2.9) ~ (1.2.13).

$$\int_0^{y_i}\left\{1-\prod_{\lambda=1}^{i-1}(1-K_\lambda)\right\}f\,dx\simeq\int_0^{y_i}K_{i-1}f\,dx$$

$$\simeq\sum_{n\geqslant 0}(-)^n E_{n+1,i}\left(\frac{d^n f}{dx^n}\right)_i\left(\frac{y_{i-1}}{\log R_{i-1}}\right)^{n+1}$$

where

$$E_{n,i}=\sum(-)^\nu\frac{1}{(\nu+1)^n}\left(\frac{1}{R_{i-1}}\right)^{\nu+1}$$

$$\times\exp[(\nu+1)\,\sigma_{i-1}\,y_i],$$

$$y_i=\frac{\log R_i}{\sigma_i}$$

$$\int_0^{y_i}K_i\prod_{\lambda=1}^{i-1}(1-K_\lambda)f\,dx-\int_y^{\infty}(1-K_i)\prod_{\lambda=1}^{i-1}(1-K_\lambda)fdx$$

$$\simeq-2\sum_\nu C_{2\nu+2}\frac{d^{2\nu+1}}{dx^{2\nu+1}}\left\{\prod_{\lambda=1}^{i-1}(1-K_\lambda)f\right\}\left(\frac{y_i}{\log R_i}\right)^{2\nu+2}$$

By using these relations, Eq. (2.3.10) can be reduced to

$$\int_{x_i}^{y_i}fdx\simeq\left[E_{1,i}\left(\frac{y_{i-1}}{\log R_{i-1}}\right)-2C_2\frac{d}{dx}\left\{\prod_{\lambda=1}^{i-1}(1-K_\lambda)\right\}_i\times\left(\frac{y_i}{\log R_i}\right)^2\right]f(x_i)+\left[-E_{2,i}\left(\frac{y_{i-1}}{\log R_{i-1}}\right)^2-2C_2\times\prod_{\lambda=1}^{i-1}(1-K_\lambda)_i\left(\frac{y_i}{\log R_i}\right)^2\right]\left(\frac{df}{dx}\right)_i+\cdots\cdots. \qquad (2.3.11)$$

The restrictions for $f(x)$, such as (1. I) and (1. II) in Section 1, should be also satisfied by the above equations.

As x_i is defined for $f(x)$, so x_i' is defined for $x^\alpha f(x)$.

$$\int_0^{x_i'} x^\alpha f(x)dx=\int_0^\infty \prod_{\lambda=1}^{i} \{1-K_\lambda(x)\}x^\alpha f(x)dx, \quad (2.3.12)$$

$$\int_{x_i'}^{x_{i-1}'} x^\alpha f(x)dx=M_i, \quad (2.3.13)$$

$$\begin{aligned}\int_{x_i}^{y_i} x^\alpha f(x)dx\simeq&\Big[E_{1,i}\Big(\frac{y_{i-1}}{\log R_{i-1}}\Big)-2C_2\frac{d}{dx}\\ &\times\Big\{\prod_{\lambda=1}^{i-1}(1-K_\lambda)\Big\}_i\Big(\frac{y_i}{\log R_i}\Big)^2\Big]x_i^\alpha f(x_i)\\ &+\Big[-E_{2,i}\Big(\frac{y_{i-1}}{\log R_{i-1}}\Big)^2-2C_2\prod_{\lambda=1}^{i-1}(1-K_\lambda)_i\\ &\times\Big(\frac{y_i}{\log R_i}\Big)^2\Big]\Big(\frac{d(x^\alpha f)}{dx}\Big)_i+\cdots\cdots. \quad (2.3.14)\end{aligned}$$

The right-hand side of this equation is reduced to

$$\begin{aligned}&\simeq\sum(-)^n E_{n+1,i}\Big(\frac{d^n(x^\alpha f)}{dx^n}\Big)_i\Big(\frac{y_{i-1}}{\log R_{i-1}}\Big)^{n+1}\\ &-2\sum C_{2\nu-2}\frac{d^{2\nu+1}}{dx^{2\nu+1}}\Big\{\prod_{\lambda=1}^{i-1}(1-K_\lambda)x^\alpha f\Big\}\\ &\times\Big(\frac{y_i}{\log R_i}\Big)^{2\nu+2}\simeq y_i^\alpha f(y_i)\Big[E_{1.i}\frac{y_{i-1}}{\log R_{i-1}}\\ &-2C_2\frac{d}{dx}\Big\{\prod_{\lambda=1}^{i-1}(1-K_\lambda)\Big\}_i\Big(\frac{y_i}{\log R_i}\Big)^2\\ &+y_i^\alpha\Big(\frac{df}{dx}\Big)_i\Big[-E_{2,i}\Big(\frac{y_{i-1}}{\log R_{i-1}}\Big)^2\\ &-2C_2\prod_{\lambda=1}^{i-1}(1-K_\lambda)_i\Big(\frac{y_i}{\log R_i}\Big)^2\\ &+\alpha y_i^{\alpha-1}f(y_i)\Big[-E_{2,i}\Big(\frac{y_{i-1}}{\log R_{i-1}}\Big)^2\\ &-2C_2\prod_{\lambda=1}^{i=1}(1-K_\lambda)_i\Big(\frac{y_i}{\log R_i}\Big)^2\Big]+\cdots\cdots,\end{aligned} \quad (2.3.15)$$

$$\begin{aligned}&\simeq\int_{x_i}^{y_i} x^\alpha f(x)dx-\alpha y_{i-1}^{\alpha-1}f(y_i)\Big[E_{2,i}\Big(\frac{y_{i-1}}{\log R_{i-1}}\Big)^2\\ &+2C_2\prod_{\lambda=1}^{i-1}(1-K_\lambda)_i\Big(\frac{y_i}{\log R_i}\Big)^2+\cdots\cdots.\end{aligned} \quad (2.3.16)$$

Thus, from Eq. (2.1.7), Δ_i is approximately equal to

$$\begin{aligned}\Delta_i\simeq&\alpha y_{i-1}^{\alpha-1}f(y_{i-1})\Big[E_{2,i-1}\Big(\frac{y_{i-2}}{\log R_{i-2}}\Big)^2\\ &+2C_2\prod_{\lambda=1}^{i-2}(1-K_\lambda)_{i-1}\Big(\frac{y_{i-1}}{\log R_{i-1}}\Big)^2\Big]\\ &-\alpha y_i^{\alpha-1}f(y_i)\Big[E_{2,i}\Big(\frac{y_{i-1}}{\log R_{i-1}}\Big)^2\\ &+2C_2\prod_{\lambda=1}^{i-1}(1-K_\lambda)_i\Big(\frac{y_i}{\log R_i}\Big)^2\Big].\end{aligned} \quad (2.3.17)$$

The mean value theorem of the integral gives

$$\int_{x_i}^{x_{i-1}} x^\alpha f(x)dx=(x_{i-1}-x_i)\xi_i^\alpha f(\xi_i). \quad (2.3.18)$$

$$x_i\leq\xi_i\leq x_{i-1}$$

In a rough approximation, condition (2. I) may be replaced by the following inequality (2. II) by the use of Eqs. (2.3.17) and (2.3.18). The only one squarely bracketed term in the right side of Eq. (2.3.17) was used in its derivation.

$$\{E_{2,i-1}+2C_2(1-K_{i-2})_{i-1}\}\frac{\alpha}{(\log R)^2}\ll\frac{x_{i-1}-x_i}{x_{i-1}} \quad (2.\ \mathrm{II})$$

R: smallest R among R_i, R_{i-1} and R_{i-2}.

The above derivation is not mathematically strict, but as it seems nearly hopeless to expect a more rigorous solution; we must content ourselvelves with the use of (2. II). It seems, however, that there are no appreciable errors in this approximation. This point will be discussed again in the next paragraph. The magnitude of $E_{2,i-1}+2C_2(1-K_{i-2})_{i-1}$ never exceeds $\log 2+\pi^2/6$.

The final result is

$$\left(\log 2+\frac{\pi^2}{6}\right)\frac{\alpha}{(\log R)^2}\ll\frac{x_{i-1}-x_i}{x_{i-1}}. \qquad (2.\text{III})$$

Here, we are again confronted with a difficult question, i.e. how to estimate x_i. We are obliged to content ourselves with a somewhat non-rigorous estimate as follows:

$$x_i=\frac{W_{i+1}m_i^{1/\alpha}+W_i m_{i-1}^{1/\alpha}}{W_{i+1}+W_i}. \qquad (2.3.19)$$

2.4. Discussion

From the above discussions, it is concluded that there are two important factors in fractionation, one being to make the amount of the fractionated polymer as small as possible and the other being to satisfy condition (2.III) in each fractionation step.

In deriving (2.II) from Eq. (2.3.18), it is assumed implicitly that the differences between $f(y_{i-1})$ and $f(y_i)$ or among y_{i-2}, y_{i-1}, and y_i are not very large. These undesirable assumptions, however, are not very important, because we use only one of the two squarely bracketed terms in Eq. (2.3.17).

Perhaps the most problematical point in our treatment is the replacement of x_i by Eq. (2.3.19).

Conclusion

By the use of Flory's theory, in Section 1, equations concerning Spencer's method were formulated to obtain the molecular weight distribution of high polymers. The author proposed to regard these equations as a set of simultaneous integral equations with respect to the distribution $f(x)$ and a function of solution composition $\sigma(s)$. Although these equations required knowledge of the ratio of the volume of the supernatant phase to that of the precipitated phase, an approximate solution was presented in this paper. This method may be considered to be a modification of Spencer's method which is more reliable than the original. The author's theory is confirmed by numerical calculations made by himself and by other investigator.$^{(6)}$

In Section 2, the author discussed the reliability of the distribution obtained by the successive precipitation method (SPM). Considering the effects of the overlap of distribution of each fraction, the author proposed a criterion for the correctness of the distribution obtained by SPM. By the use of Flory's theory, the proposed criterion was written down in a closed mathematical from as Eq. (2.III). In the derivation of this relation, we were obliged to use a somewhat ambiguous approximation which is not mathematically strict but seems not to be so erroneous as to mislead.

Appendix

Proof of the inequality.

$$E_{2,i-1}+2C_2(1-K_{i-2})_{i-1}<1+\pi^2/6.$$

Proof:

$$E_{2,i-1}=\sum_{\nu\geqslant 0}(-)^{\nu}\frac{1}{(\nu+1)^2}\left(\frac{1}{R_{i-1}}\right)^{\nu+1}\exp[(\nu$$

$$-1)\sigma_{i-1}y_i]=\sum_{\nu\geqslant 0}(-)^{\nu}\frac{1}{(\nu+1)^2}a^{\nu+1} \qquad (\text{A}.1)$$

where

$$a=(R_{i-1})^{y_i/y_{i-1}-1}.$$

The last series converges uniformly for $0<a<1$. The series obtained by term-by-term differentiation also converges uniformly for $0\leqslant a\leqslant 1$. Therefore the following differential equation holds.

$$\frac{dE_{2,i-1}}{da}=1-\frac{a}{2}+\frac{a^2}{3}-\frac{a^3}{4}+\cdots\cdots$$

$$=\frac{1}{a}\log(1+a), \qquad (\text{A}.2)$$

$$E_{2,i-1}=\int_0^{\alpha}\frac{1}{a}\log(1+a)da. \qquad (\text{A}.3)$$

Eq. (A.3) cannot be integrated in a compact form but Eq. (A.2) tells us that $E_{2,i-1}$ is increases monotonically for $0 \leq a \leq 1$.

$$E_{2,i-1}(a) < E_{2,i-1}(1).$$

A theorem about alternative series tells us that

$$E_{2,i-1}(1) = \frac{1}{2}\left\{1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots\cdots\right\} < 1$$

$$E_{2,i-1}(a) < 1$$

and the following relation is almost evident.

$$2C_2(1-K_{i-2})_{i-1} < 2C_2 = \pi^2/6$$

Finally,

$$E_{2,i-1} + 2C_2(1-K_{i-2})_{i-1} < 1 + \pi^2/6.$$

References

(1) F. W. Billmeyer & W. H. Stockmayer; "Method of Measuring Molecular Weight Distribution," *J. Polymer Sci.*, **5**, p. 121 (1950).

(2) R. S. Spencer; "Determination of Molecular Weight Distributions," *J. Polymer Sci.*, **3**, p. 606 (1948).

(3) G. V. Schulz; "Die Verteilungsfunktionen Polymolekularer Stoffe und Ihre Ermittelung durch Zerlegung in Fraktionen," *Zeit. Phys. Chem.*, **B**, **47**, p. 155 (1940).

(4) E. V. Syre; "The Distribution of Polymer Fractions," *J. Polymer Sci.*, **10**, p, 175 (1953).

(5) M. Katumoto & Y. Oyanagi; "A Note of the Successive Fractionation Method," *Chemistry of High Polymers*, **11**, p. 7 (1954).

(6) M. Matumoto; "Theoretical Considerations on the Spencer's Fractionation Method," *Chemistry of High Polymers*, **11**, p. 182 (1954).

(7) P. J. Flory; "*Principles of Polymer Chemistry.*" p. 559, 1953, Cornell Univ. Press.

(8) R. S. Scott; "Thermodynamic of High Polymer Solutions: II The Solubility and Fractionation of a Polymer of Heterogeneous Distributions," *J. Chem. Phys.*, **13**, p. 178 (1945).

* * * *

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 1-2, JANUARY-FEBRUARY, 1961

U.D.C. 621.374 : 621.375.3

Core Matrix Driver Employing Magnetic Amplifiers*

Junji YAMATO and *Yasunobu SUZUKI*

A high frequency magnetic amplifier with high efficiency, high S/N, and high output power has been developed to drive large scale core memory matrixes in parametron devices.

Introduction

In many electronic digital computers and electronic switching systems employing parametron logical devices, memory core matrixes are provided for storing digital information. In such cases, the so-called two frequency alternating current writing method is widely used, in which a sinusoidal current of $f/2$ is superposed on a sinusoidal current of frequency f to form a dissymmetrical current, which is used for writing on cores.

In a large scale electronic digital computer or switching system, a large memory capacity is required, thus necessitating a large scale memory core matrix and, as a result, large power output and a high S/N are required of the power source which drives the memory core matrix.

Where, for instance, a memory matrix of 10^4 bits (10^2 columns $\times$ 10^2 rows) is to be driven, a peak power output of 10 Watts is required. It is, however, apparent that the use of 200 power amplifiers is not desirable from the view-points of power loss and economy. On the contrary, another method may be feasible, whereby power is supplied from a common source to the terminal selected from among the many input terminals by means of an electronic switch, enabiling more economical operation. Needless to say, the electronic switch used for this purpose must have long life, low cost, high efficiency, high reliability, no deterioration, small size, and high switching ratio.

From these points of view, the writers concentrated on a high frequency magnetic amplifier and developed a magnetic core driver whose S/N ratio may be raised through the application of balanced bridge theory and from which an output of about 10 Watts may be obtained easily using a ring-shaped ferrite core with an outer diameter of 8 mm.

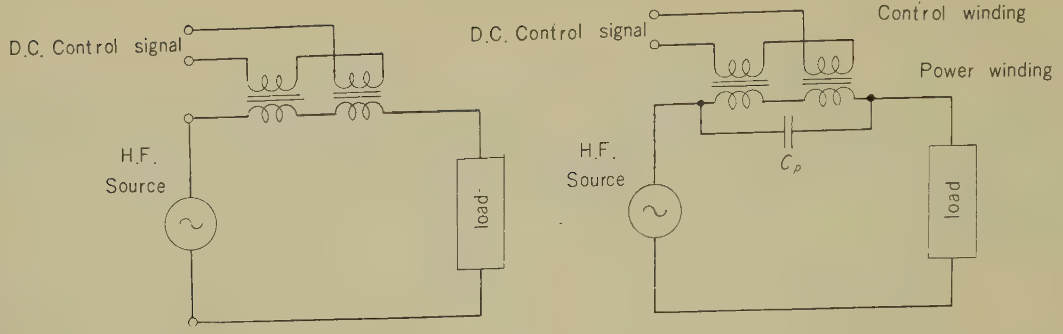
1. Improvement of the Switching Ratio

Since a magnetic amplifier is essentially a variable impedance element in an alternating current circuit, it may be used effectively for performing ON-OFF functions for high frequency alternating currents. Ferrite cores are used almost exclusively in high-frequency magnetic amplifiers, but the switching ratio of ferrite cores is quite low compared with the ratio obtainable with low frequency switches using the usual types of core materials; therefore the improved circuit shown in Fig. 1 must be used. Fig. 1, (a) shows the circuit generally used, while (b) shows a circuit which is parallel tuned at the zero control signal condition, thereby raising the switching ratio by about 10 dB compared with circuit (a); and (c) shows a circuit which is series tuned by C_s at the driving condition,

---

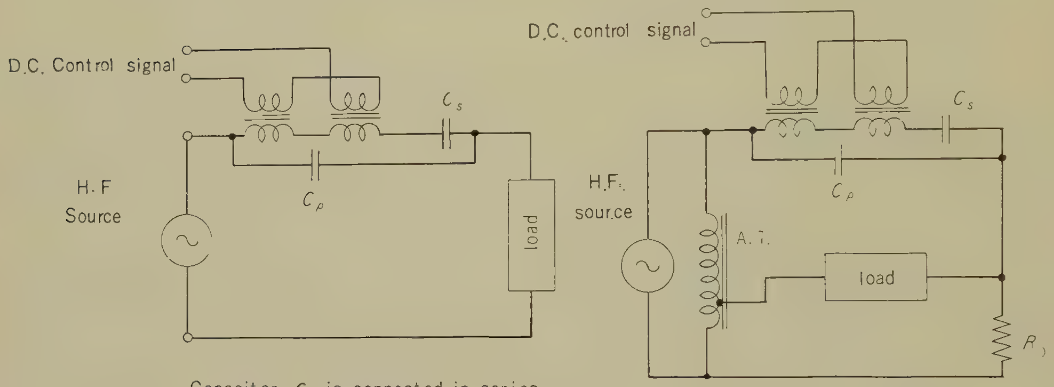
* MS in Japanese received by the Electrical Communication Laboratory, Aug. 25, 1960. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T., Vol. 9, No. 5, May, pp. 537–563, 1960.

† Switching Research Section.



(a) Fundamental circuit of no-feedback type series-connected magnetic amplifier

(b) Modified fundamental circuit of the magnetic amplifier with capacitor which parallel turns the inductance of the power winding when no d.c. current is fed to the control winding.



Capacitor C_s is connected in series with the power winding. Capacitor C_s series tunes.

By constructing a bridge using an autotransformer and resistance R_0, leakage current from the magnetic amplifier is greatly reduced.

Fig. 1—Improvement of magnetic switch circuit.

thereby further improving the switching ratio by 10–12 dB compared with circuit (b).

A switching ratio of about 40 dB may be obtained quite easily with the circuit shown in (c). For driving a large-scale core matrix, however, a switching ratio of 40 dB is not sufficient. Fig. 1, (d) show a circuit in which a bridge is formed with a magnetic amplifier, a balancing resistor, and an autotransformer, thereby cancelling the leak at the zero control signal condition so that, in theory, infinite S/N is possible. But an S/N of about 50 dB is the practical limit owing to the production spread of various parts used for many electronic switches and to the second harmonics induced from the cores used for the magnetic amplifiers.

2. Magnetic Amplifier Control Circuit

Address selection signal of the experimentally made core memory matrix will be given by parametron signals consist of "one out of *n* codes" system. Fig. 2 shows, as an example, the circuit arrangement of a 20 row driver. Signals S_1, S_2,····S_5, and S_1', S_2',····S_4' are all parametron signals (carrier frequency 1 Mc, modulating frequency 10–25 kc, duty ratio 50%), and the information is expressed as "1" when it is in-phase. and as "0" when it is 180° out of phase with the standard phase and phase-voltage con-

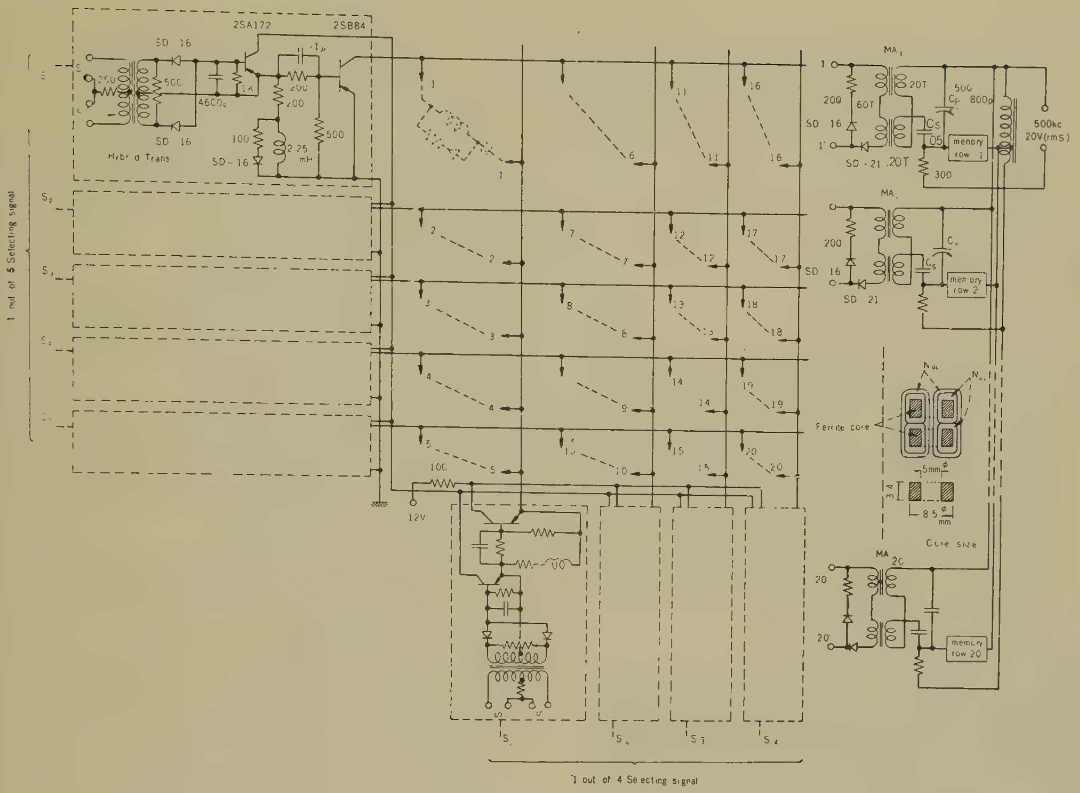


Fig. 2—Circuit arrangement of 20 row driver with an output of 14 Watts each.

version is made by comparsion with the standard phase.

The two-stage transistor amplifier shown within the dashed lines works like an ON-OFF switch in a direct current circuit, so that when a switch in one row and a switch in one column are ON, current will flow only through the control winding of the magnetic amplifier at the crossing point, thereby turning the corresponding magnetic switch ON, thus supplying sufficient writing current to the driving loop of core matrix connected to the switch.

The input power obtained from parametrons S_n, S_n', S_0, and S_0' is about 2.5 mWatts each; which exceeds the 1–2 mWatts driving power required by the 2SA172 transistors. Transistor 2SA172 will switch ON a maximum current of 200 mA, including the charging current for the 0.1 μfd speed-up condencer, and a static current of 120 mA, thus supplying sufficient driving current to power transistor 2SB84. Transistor 2SB84 ($f\alpha \doteqdot 400$ KC: equivalent to RCA 2N301A) able to switch a current of 1.2 Amp within 3μ sec; and within 7–8 μ sec when the time constant of the control winding of the magnetic amplifier is included.

In order to shorten the time required for turn on and turn off, a 0.1 μfd speed-up condenser and a 2.25 μH reverse base-driving choke coil are used, the former is effective in shorten—the "turn-on" time and the latter inversely shortening the "turn-off" time.

In the OFF condition of the magnetic amplifier, a tuned impedance of more than 1 kΩ can be obtained, while in the ON condition, a series tuned impedance of about 1Ω can be obtained. Therefore, for a load resistance of 14Ω, an ON OFF ratio of 1014/15, i.e. an S/N of about 37 dB may be obtained even with the circuit shown in

Fig. 1 (c). If, in this case, a bridge connection is formed as shown in Fig. 1 (d), an S/N of at least 50 dB may be secured. The driving circuit impedance of memory-core matrix will be transformed into 14 Ω through a matching transformer.

3. Design Consideration

a) Magnetic Amplifier

Almost all magnetic amplifiers used in the low frequency band are of the self-saturating or doubler type. However, the case of magnetic amplifiers in the high frequency band, feedback types are not used since such types is suffer from extreme deterioration of the characteristics of the core materials and feedback rectifiers, so that no great difference can be seen in characteristics between feedback and non feedback types.

The following general relation exists between the core volume and the output power of magnetic amplifiers.

$$\mathrm{Vol}=\frac{w_0\times 10^8}{\frac{2\pi}{\sqrt{2}}\cdot f\cdot B_m\cdot\frac{A}{\mathrm{cm}}\cdot K_1\cdot K_2}\ (\mathrm{cm}^3) \qquad (1)$$

where

w_0—output power (Watts)
f—source frequency (C/S)
B_m—maximum flux density (Gauss)
K_1—maximum output voltage/source voltage
K_2—maximum output current/ideally saturated current.

By introducing output=14 Watts, $f\fallingdotseq 500$ kc, $B_m=400$ Gauss, $A/\mathrm{cm}=10$ and K_1, $K_2=0.9$ into Eq. (1), a core volume of about 0.1 cm^3 can be obtained. Therefore, a core size of 5 mm inner diameter, 8.5 mm outer diameter and 3.4 mm ring width be quite satisfactory.

As against 24 A.C. ampere turns, 60 control ampere turns are required in order to have the core fully saturated, so that 60 turns on the control winding and 1.2 Amp. of control current were decided upon (a 20% margin was allowed.).

b) Optimum Load Impedance

The ON impedance of the magnetic amplifier will be, as mentioned above, lowered to about 1 Ω due to series tuning, so that, in order to introduce maximum output, the load impedance should also be fixed at 1 Ω. In this case, however, the output voltage in the ON condition will be 1/2 of that for the no-load condition and a large variation will be seen in output voltages due to the variations between electronic switches.

Assuming now that the circuit reactance is X and that the load resistance is R when the magnetic amplifier is ON. If the balancing resistance of the bridge circuit is neglected the following relation will exist between the variation of reactance X and the variation of the output voltage.

$$\frac{\Delta IR}{IR}=-\left(\frac{X}{z}\right)^2\left(\frac{\Delta X}{X}\right)=\frac{\Delta V_0}{V_0} \qquad (2)$$

where

$$z=(X^2+R^2)^{1/2}.$$

In case, for example, the reactance in the ON condition deviates from the reference value by ±90% (0.1 Ω∼1.9 Ω), $R/X=6.6$ may be obtained from Eq. (2) in order to make the variation in output voltage stay within ±2%. Namely, it is apparent that a load impedance of at least 6.6 Ω is required. Also, in order to insure that the phase variation in the ON condition is within ±5°, the load resistance must be greater than 14 Ω. Thus, the level of driving voltage of the memory core matrix may be stabilized to a high degree if careful consideration is given to safety factors when designing the magnetic amplifier.

c) Balancing Resistance *R*

The determination of the resistance of the balancing resistor R_0 is closely connected with the leak level in the OFF condition. Suppose for instance, that the parallel turned

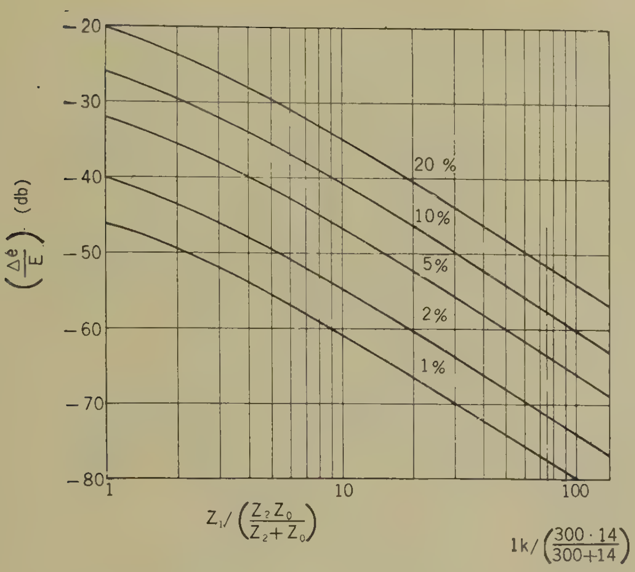


Fig. 3—Leakage output vs. impedance ratio.

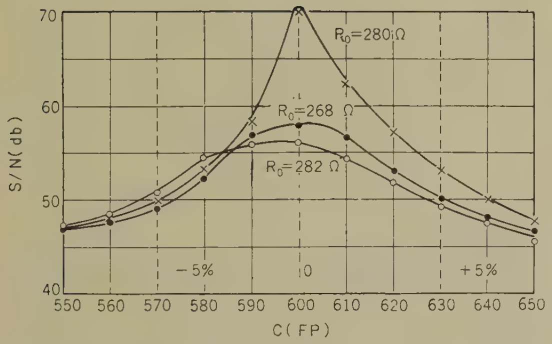
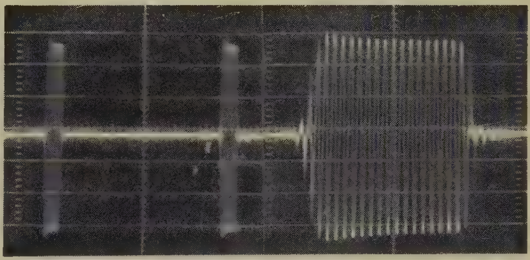
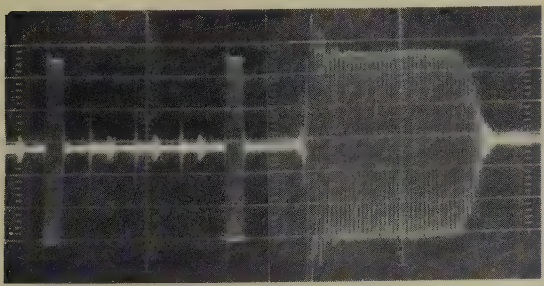


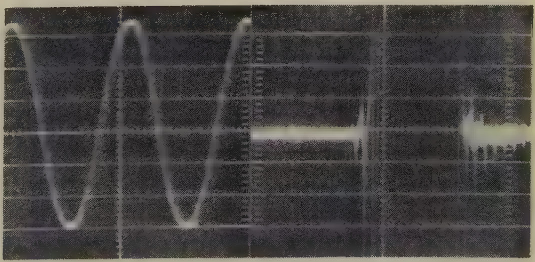
Fig. 4 (A)—S/N vs. balancing resistor R_0 and parallel tuning condenser C_p (25°C).



(ii)—Output signal waveform (500 kc) (left). Same signal but time scale expanded 10 times (right).



(i)—Input controlling signal (1 Mc) (left). Same signal but time scale expanded 10 times (right).



(iii)—Waveform of signal in output with time scale expanded 100 times (left). Waveform of noise in output with voltage scale expanded 10 times (right).

Fig. 4 (B)—Input and output waveforms of core matrix driver.

impedance is Z_1 in the OFF condition of the electronic switch, R_0 is Z_2 and Z_1 to be accompanied with variance ΔZ_1, no load leak voltage δl_0 caused by unbalance of bridge may be represented as

$$\delta l_0 = E\left[\left(1+\frac{z_1}{z_2}\right)^{-1}\left(\frac{\Delta z_1}{z_1}\right)-\left(1+\frac{z_2}{z_1}\right)^{-1}\left(\frac{\Delta z_1}{z_1}\right)^2\right. \tag{3}$$

where

E = maximum output voltage.

Therefore, the relation between the leak voltage and the impedance ratio in the loaded condition may be plotted as shown in Fig. 3. As seen from this figure, a leak level of −50 dB may be maintained with the circuit constants shown in Fig. 2 if the imdedance variation of Z_1 is within 20%.

Fig. 4 (A) shows the measured values of S/N against the variation of balancing resistance R_0 and parallel tuning condenser C_p, while Fig. 4 (B) shows the input signal waveform Fig. 4 (B), i and output voltage waveform Fig. 4 (B), ii of the magnetic amplifier. These photographs show the waveforms in case where the same row was driven once for every four modulating cycles.

4. Output Waveform, Efficiency, and Other Characteristics

a) Output Waveform Distortion

The output waveform is, as shown in Fig. 4 (B), approximately a sine wave. Measurement of the harmonic components showed that there was a second harmonic component of approximately −40 dB due to hysteresis in the core of the magnetic amplifier and a third harmonic component of about the same value caused by the nonlinearity of the core. As a matter of course, the second harmonic depends on the source voltage, having no bearing on control ampere turns, and the third harmonic diminishes as the control ampere turns increase.

b) No Load Loss of Magnetic Amplifier

The power loss caused by the core and balancing resistance of the magnetic amplifier in the unexcited condition is about 0.3 Watts per bridge. Therefore, supposing for instance that an output of 14 Watts was dissipated by one driver out of 100 row drivers and that 0.3 Watts was dessipated by the remaining 99 drivers, a high frequency source supply of about 44 Watts is required, and the efficiency is 32%. If in this case, all the magnetic amplifiers are replaced by vacuum tube amplifiers, the apparent efficiency will be below 1%.

c) Temperature Characteristics

The variation of the characteristics of the ferrite core and the transistor under which ambient temperature variation can not be disregarded. Temperature coefficient of the permeability of ferrite core, are about 0.5 %/°C at the temperature of approximately 20°C. Therefore, as ambient temperature varies by ±10°C, inductance varies by ±5%.

Since, this corresponds to a variation approximately ±5% of the parallel tuning condenser C_p shown in Fig. 4 (A), it can be understood that an S/N of −50 dB can be maintained for a ±10°C variation of ambient temperature, if adjustment is made so as to bring about the optimum conditions at the temperature of 25°C.

In the transistor amplifier, variation occurs mainly in I_{co} and in the level between base and emitter. As regards I_{co}, however, slight Quiescence current is flowing through the control winding of the magnetic amplifier at the OFF condition, and this has a tendency to reduce inductance as against the rise in temperature, thereby somewhat compensating for the increase of inductance due to the rise of temperature of the ferrite core.

d) Output Voltage Stability

Since, as mentioned before, the load impedance is selected to be more than 10 times as large as the reactance of the ac winding in the ON condition, the variation of output voltage due to load variation, transistor D.C. source voltage fluctuation, and input signal level variation is very slight. For instance, for a D.C. sourse voltage variation of ±10% and an input signal variation of ±10%, the reactance in the ON conditions is on the average, only 1Ω±21%, so that even if the load resistance also varies, by ±10%, the variation of output voltagem aintaining within ±1.4%.

e) Characteristics Required of High Frequency Power Source

As the output power is distributed from a common source by means of switching, the efficiency will be lowered unless the power source impedance is kept much lower than the load impedance. The high frequency power source, therefore, should be designed so as to have a sufficiently low output impedance by using a cathode follower, servomechanism or the like.

Conclusion

A core matrix driver employing as an electronic switch a magnetic amplifier from which a high switching ratio can be obtained has been described. What characterized this system is that by the combination of power transistors and magnetic amplifiers, high frequency power of 500 kc–2 Mc with an output of 10–20 Watts can easily be distributed to many driving terminals of a core matrix with the improved reliability, S/N, and output stability for which strenuous efforts have been made. The volume occupied by each driving unit is about 1×1×2 inches, exclusive of the common power supply.

Although in this paper the explanations are limited only to the core matrix driver; needless to say, this system may also be utilized for driving a magnetic core converter(3) or a sort of permanent memory so-called metal card memory for parametron device.

References

(1) Eiichi GOTO, "On the Application of Parametrically Exited Non-linear Resonators," *J.I.E.C.E. of Japan.* Vol. 38, No. 10, pp. 770, Oct., 1955.

(2) Hidetoshi Takahashi and Eiichi Goto' "Core Memory for Parametron Device," *Computer Conference of I.E.C.E. of Japan*, Feb., 1956.

(3) Koshiro Hanawa and Kikunobu Kusunoki, "Signal Converter by Magnetic Core for Parametron Device," *Reports of the E.C.L.*, NTT, Vol. 7, No. 2, pp, 25–31, Feb., 1959.

* * * *

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 1-2, JANUARY-FEBRUARY, 1961

U.D.C. 666.1/2:539.9

A Consideration on Relation between Mechanical Behavior of Glasses and Their Internal Structure*

Muneo WATANABE† and *Taro MORIYA*‡

The relation between the mechanical properties of glasses and their internal structures is theoretically considered. First, experimental results of the time dependence of viscosity, visco-elasticity, and strength of glasses are clearly indicated. Then the "glassy microphase" theory, a previously published hypothesis on the internal structure of glass, is applied to the results. The theory and the experimental data are in good agreement.

Introduction

Watanabe, one of the present authors, has reported in many papers(1),(2),(3),(4) concerning the experimental results of experiment on the mechanical properties of glasses such as viscosity, visco-elasticity, and strength. In other papers,(5),(6) he also reported on his studies of the internal structure of glasses by the electron microscope.

The purpose of this paper is to consider theoretically the mechanical properties of glasses in the light of the "glassy microphase" theory.

This paper consists of three parts: first, an explanation of the experimental results on the internal structure of glasses; second, a description of experimental results on the mechanical properties of glasses; and third, a theoretical consideration on the relation between the mechanical properties of glasses and the internal structure of glasses.

1. The Internal Structure of Glasses.

The internal structure of glasses deduced from electron microscope studies(5) can be summarized as follows:

(a) Glass which is seemingly transparent and homogeneous, is in fact hetrogeneous in the sense of micro-structure. This heterogeneity consists matrix and micro-structures which the authors called "glassy-microphases".

(b) These glassy microphases have certain settled average chemical compositions, such as sodium borate and silica rich in borosilicate glasses.

(c) These glassy microphases are not crystal, as usually conceived; but amorphous, so to speak.

(d) The size of the glassy microphase grows with increase in heat-treating time at constant temperature at a given point within the transformation range.

(e) The size of the glassy microphase varies with different heat-treating temperatures for a constant heat treating time. Its size in a heat-treated specimen is bigger than that in a quenched specimen.

These five experimental results are schematically expressed in Fig. 1, which shows how the size and shape of the glassy microphases of two different chemical compositions changes with heat-treating time for constant heat-treating temperature. A similar pattern of change takes place if the temperature varies and the heat-treating time is constant.

---

* MS in Japanese received by the Electrical Communication Laboratory, Sept. 19, 1960.

† Ferroelectric Materials Research Section.
‡ The Tokyo Institute of Technology.

The basic principle of a schematic representation such as Fig. 1 has been reported by Moriya.(7),(8) Watanabe proved its validity experimentally, and gave it a quantitative expression in Fig 2.

Fig. 2 illustrates how the size of the SiO_2 glassy microphases change with heat-treating time for two different temperatures. The term "width" in Fig. 2 stands for the diameter of the spherical glassy microphases when they are isolated from one another, as shown in (A) and (B) of Fig. 1. When they combine and form skeletons as shown in (C) of Fig. 1, the "width" stands for the shorter width of the skeleton.

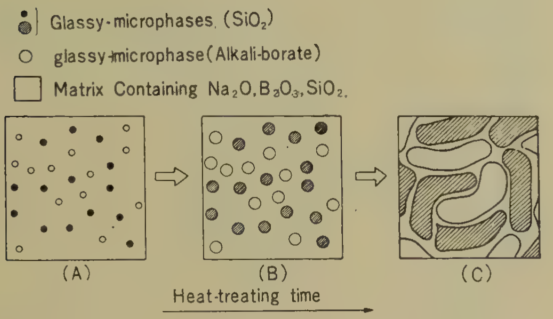

Glassy-microphases (SiO₂)
glassy-microphase (Alkali-borate)
Matrix Containing Na₂O, B₂O₃, SiO₂.
(A)
(B)
(C)
Heat-treating time

Fig. 1—A schematic representation of the change of the internal structure of a glass.

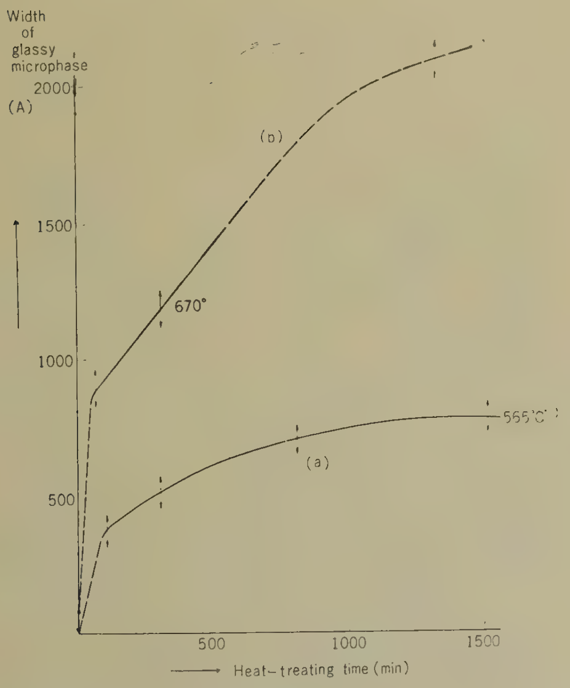


Fig. 2—Growth of the "Glassy microphase" (SiO_3) with heat-treating time and temperature.

The shorter width of a skeleton has the same "width" as an isolated spherical microphase produced under the same heat-treating conditions, therefore there is no theoretical difficulty in defining "width".

It may not necessarily follow that glassy microphase patterns identical to those shown in Fig. 1 appear in all kinds of glass. But the authors consider that regardless of difference in size, shape, and chemical composition, glassy microphases exist in every kind of glass. This point is elaborated on in Appendix 1.

In this paper, the "glassy microphase" theory is used as a basis of the explanation of the experimental results of the mechanical properties of glasses.

2. A Consideration on the Relation between the Strength of Glasses and Their Internal Structure.

The strength of glass is one of its mechanical properties, and is indicated by its fracture under a maximum stress. Concerning this property of glass, i.e. strength, there is Griffith's theory(9) which has been confirmed by many papers.

On the other hand, the "fatigue" phenomenon pointed out by Preston(10) has become well-defined since R.E. Mould(11) and R.D. Southwick found the universal fatigue curve.

In this section, the authors consider the relation between the internal structure of glasses and their strengths, for both abraded and unabraded specimens, and for those with and without the fatigue phenomenon.

2.1. Symbols used in This Section

σ=strength (general)

σ_n=strength of unaged specimens tested in liquid nitrogen

σ_n'=strength of aged specimens tested in liquid nitrogen

σ_w'=strength of aged specimens tested in water

C=depth of cracks (general)

C_0=depth of abrasions artificially made on the surface of unaged specimens

C_0'=depth of abrasions artificially made on the surface of specimens aged over water

C_1=depth of latent flaw due to the boundary size of the glassy microphases in freshly drawn specimens (i.e., specimens which have not been heat treated.)

C_2=depth of latent flaw due to the boundary size of the glassy microphases in heat-treated specimens

C_3=depth of latent flaw produced by a physicochemical reaction which supposedly takes place between water and glassy microphases under stress. This physicochemical reaction is considered as a cause of the fatigue phenomenon.

E=Young's Modulus

α=Surface tension

2.2. Strength of Glasses without the Fatigue Phenomenon

Griffith's formula,⁽⁹⁾ which is given below, can be use for the strength of glasses without fatigue.

$$\sigma=k\sqrt{\frac{E\cdot\alpha}{C}} \qquad (1)$$

where k is constant ($\fallingdotseq 1$).

2.2.1. Griffith's Flaw

As reported in Watanabe's previous paper,⁽³⁾ the specimens were prepared with extreme care so as to avoid any possible damage before their strength was measured. The values of the strength of the unabraded specimens tested in liquid nitrogen were still much lower (150,000–180,000 p.s.i.) than the theoretical values.

This low strength is ascribed to Griffith's flaws latent in the unabraded specimens. But why do the specimens which are neither abraded nor touched have such a latent flaw?

To answer this question, the authors assumed that the latent Griffith's flaw is attributed itself the boundary line of the glassy microphases in the glass specimens. This assumption is illustrated in Fig. 3. The size of the glassy microphase in a quenched specimen is small, whereas that in a heat-treated specimen is large, as shown in (A) and (B) of Fig. 3.

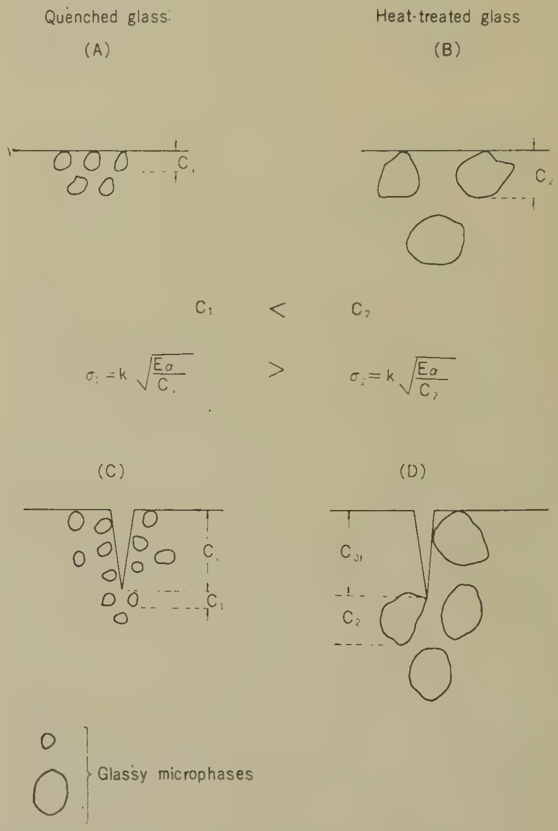


Fig. 3—Difference between the strength of quenched and heat-treated specimens.

It may well be assumed that there exists a weaker boundary layer between a glassy microphase and it's matrix part.

This layer is considered a latent flaw.

This means that each specimen has latent flaws, the depth of which corresponds to C_1 in Fig. 3 (A), or to C_2 in Fig. 3 (B). In other words, the strength σ_1, of a quenched specimen is greater than the strength σ_2 of a heat-treated specimen, which can be inferred from Eq. (1), for $C_1 < C_2$.

This assumption will successfully explain the previous experimental results(12)(13) in which the strength decreased as the heat-treatment increased. The following is an examination of the applicability of the above assumption, using Griffith's formula.

The strength value of unabraded (V–1) glass tested in liquid nitrogen was 8,400 kg/cm^2 (120,000 p.s.i.). This was a heat-treated specimen (500°C, 1 hour), and its internal structure was as illustrated in Fig. 1. From the experimental data of Spinner(14) and Parmee,(15) Young's Modulus of the specimen was estimated to be 800×10^9 dyne/cm and its surface tension to be 3×10^2 dyne/cm. Substituting these values into Eq. (1), we obtain

$$C_2 = \frac{E \cdot \alpha}{\sigma^2} = \frac{800 \times 10^9 \times 3 \times 10^2}{(8 \times 10^9)^2} = 380\ (\mathring{A}).$$

This means that the unabraded specimen must have latent flaws whose depth is about 400 Å.

On the other hand, the width of the glassy microphase of a heat-treated specimen (565°C, 1 hour) was found by observation with an electron microscope to be about 500 Å. The calculated value of the flaw depth agrees sufficiently well with the experimentally observed width of the glassy microphase, if the difference of the heat-treating temperatures is taken into consideration.

In other words, it was correct to assume that the depth of the Griffith flaw C would be closely related to the width of the glassy microphase.

2.2.2. Effect of Chemical Compositions of Glasses on Their Strength in Liquid Nitrogen

The same thought process as above can be applied to the abraded specimens illustrated in (C) and (D) of Fig. 3. As shown in these figures, each specimen is abraded to the same depth C_0, so that the total depths in these cases $C_0 + C_1$ and $C_0 + C_2$, respectively. Since $C_0 + C_1 < C_0 + C_2$, the strength of the specimen in Fig. 3 (C) is greater than that of the specimen in Fig. 3 (D).

A schematic consideration such as Fig. 3 can be used to explain the difference in the strengths of glasses of different chemical compositions.

Fig. 4* shows the Isokoms (Lines of Constant strength) of the specimens tested in the liquid nitrogen after they were abraded to the same depth C_0.

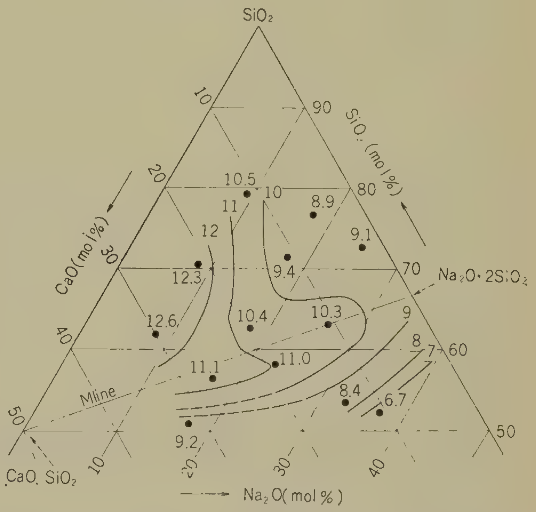


Fig. 4—The strength vs. chemical compositions of abraded and unaged specimens tested in liquid nitrogen. Strength values and contours in Kg/cm^2 × 10^2. An anomaly lies along the M line.

Apparently, Fig. 4 shows no additive relationship between the chemical composition (Na_2O, CaO, and SiO_2) and the strength σ_N.

According to Zachariasen's theory, the

---

* On Dec. 15, 1960, Watanabe received the renew data of this diagram from Preston Laboratories Inc., Butler, Pa., but main tendency is not different to this diagram and renew diagram.

strength (σ_N) should decrease monotonically in the direction of the increase in the Na_2O content of the glass, because the probability that the Na^+ion will loosen or cut the chemical bond (Si-O-Si) will increase as the amount of Na_2O increases.

This assumption, however, does not agree with the experimental results shown in Fig. 4. The isokoms make the hill of the maximum strength (i.e. anomalies) along the M line which draws from the composition $Na_2O.\ 2SiO_2$ to the $CaO.\ SiO_2$.

It should be emphasized that the chemical compositions of the both ends of the anomaly line correspond to $Na_2O.\ 2SiO_2$ and $CaO.\ SiO_2$. These are the chemical compositions corresponding to the microphases which have been theoretically suggested by Moriya.(7)

In Fig. 5, the strength values are plotted in two ternary systems according to the glassy microphase theory instead of the usual ternary system shown in Fig. 4. In Fig. 5, the chemical compositions of the glasses were recalculated from the amount of the components SiO_2, Na_2O and CaO, to the quantities of the molecular component $[SiO_2]$, $[Na_2O.\ 2SiO_2]$, $[CaO.\ SiO_2]$. No anomaly appears in the Figure and it is seen that there exists an linear relationship between the molecular compositions of the microphases and the strength.

Therefore, it can be safely said that the strength σ_N is dependent on the size or width of the glassy microphase, namely on C_2 in (C) and (D) of Fig. 3.

If this is true, the size of the glassy microphase C_2 can be calculated from Eq. (1) for various glasses. The following calculations were carried out from this point of view.

For V-1 Glass, the strength is given by the following equation:

$$\sigma_{N\cdot(V)}=k\sqrt{\frac{E_{(V)}\cdot\alpha_{(V)}}{C_0+C_{2\cdot(V)}}} \qquad (3)$$

where

$\sigma_{N\cdot(V)}$=strength of V-1 glass (unaged), tested in liquid nitrogen,

$E_{(V)}$ and $\alpha_{(V)}$=Young's Modulus and surface tension of V-1 glass,

C_0=depth of artificial abrasion,

$C_{2\cdot(V)}$=depth of latent flaw due to the boundary size of the glassy microphase. (See Fig. 3 (D)).

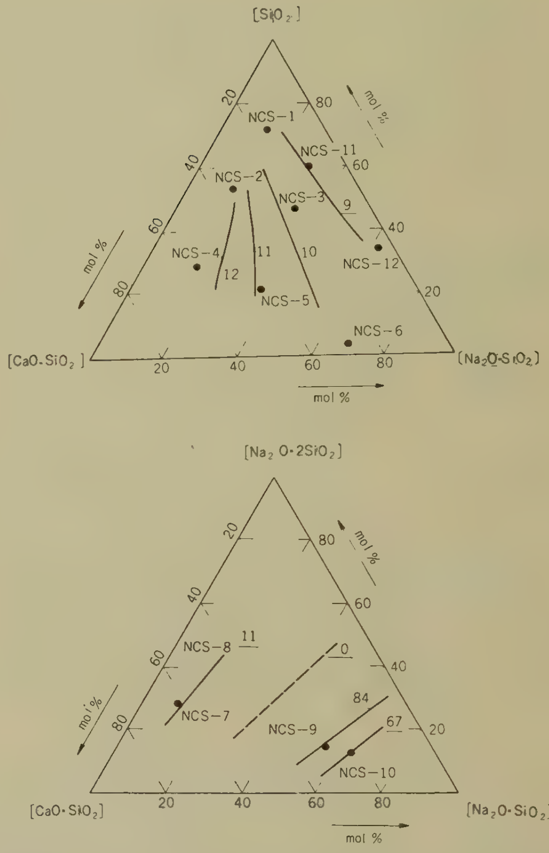


Fig. 5—Liquid nitrogen strength vs. chemical compositions on the basis of the glassy microphase theory for abraded and unaged specimens tested in liquid nitrogen. Strength values and contours in $Kg/cm^2\times10^2$. No anomaly is seen in these systems.

Similarly, for NCS-6 Glass in which the existence of glassy microphases having the composition $Na_2O.\ 2SiO_2$ is assumed, the strength is expressed as follows:

$$\sigma_{N\cdot(6)}=k\sqrt{\frac{E_{(6)}\cdot\alpha_{(6)}}{C_0+C_{2\cdot(6)}}} \qquad (4)$$

where

$\sigma_{N\cdot(6)}$=strength of NCS-6 Glass (unaged) tested in liquid nitrogen,

$E_{(6)}, \alpha_{(6)}$=Young's Modulus and surface tension of NCS-6 Glass,

C_0=depth of the artificial abrasion,

$C_{2\cdot(6)}$=depth of latent flaw due to the boundary size of the glassy microphase.

The ratio of Eq. (3) and Eq. (4) is:

$$\frac{\sigma_{N\cdot(V)}}{\sigma_{N\cdot(6)}}=\left(\frac{E_{(V)}\cdot\alpha_{(V)}}{C_0+C_{2(V)}}\right)^{\frac{1}{2}}\Big/\left(\frac{E_{(6)}\cdot\alpha_{(6)}}{C_0+C_{2(6)}}\right)^{\frac{1}{2}} \qquad (5)$$

Substituting the following experimental values in Eq. (5):

$\sigma_{N\cdot(V)}=980\ \mathrm{kg/cm^2}$
$\sigma_{N\cdot(6)}=950\ \mathrm{kg/cm^2}$
$C_0=3\times10^{-5}$ cm (See reference No. 4.)
$C_{2\cdot(V)}=5\times10^{-6}$ cm (See Fig. 2; 565°C,1 hour.)
$E_x\cdot\alpha_{(V)}\approx E_{(6)}\cdot\alpha_{(6)}\approx2\times10^{14}$ ($\mathrm{dyne^2/cm^3}$).

we obtain

$$C_{2\cdot(6)}\approx 800\ \text{Å}$$

A similar mathematical treatment was applied to the other types of glassy microphases such as $Na_2O\cdot SiO_2$, and $CaO\cdot SiO_2$.

The order of the latent flaw depths obtained are as follows:

$$[Na_2O\cdot SiO_2]>[Na_2O\cdot 2SiO_2]>[CaO\cdot SiO_2].$$

This order agrees well with the predicted size of the glassy microphases.(7)

So far, the authors have proved that there is a close relationship between the size of the glassy microphase and the strength of glasses without fatigue.

2.3. Strength of Glasses with Fatigue

2.3.1. On the Mechanism of Fatigue

The fatigue phenomenon may well be ascribed to a physicochemical reaction which takes place between the glass and water, because the fatigue phenomenon was accelerated(12)(16) by increasing the humidity during the strength test.

It is apparent that the original Griffith formula(1) can not explain the fatigue phenomenon, because in the Griffith formula, the strength is not a function of the time under stress.

Therefore, a suitable method to obtain a suitable formula is to modify the Griffith formula. In the original Griffith formula, only the flaw depth C can be modified as a function of time. Young's modulus and the surface tension cannot possibly be functions of the load duration. Even when values of Young's modulus change slightly, no big change in the value of the strength can be expected, for strength is proportional to the square root of both Young's modulus and the surface tension.

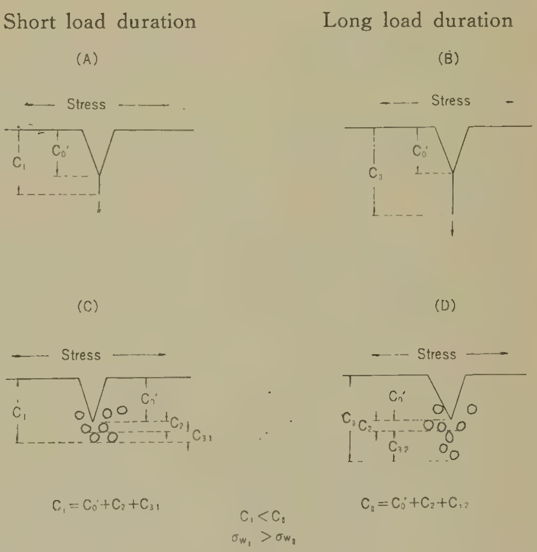


Fig. 6—A schematic representation of the mechanism of the fatigue phenomenon.

(A) and (B) of Fig. 6 illustrate the assumed mechanism of fatigue as a variable of load duration. Suppose the glass surface is artificially abraded, and the depth of the abrasion after aging is C_0'. If a tensile stress is applied in the direction of the arrow in Fig. 6, the crack depth is extended to C_{I}, owing to the reaction between the glass and water at the tip of the crack; and an explosive fracture occurs at a velocity which nearly equals the velocity of sound, when the crack becomes as deep as C_{I}.

Case (A) in Fig. 6 is for a shorter load duration than case (B). Therefore, the flaw depth C_{I} made by physicochemical reaction is shallower than C_{II}. It follows that $\sigma_{W\cdot\mathrm{I}}$, the strength tested in water for the shorter duration is greater than $\sigma_{W\cdot\mathrm{II}}$, the strength tested for the longer duration. That is, $\sigma_{W\cdot\mathrm{I}} > \sigma_{W\cdot\mathrm{II}}$.

Such an analytical explanation of the mechanism of fatigue is comparatively easy to understand. A more detailed discussion on this point will be made in a later section.

Next, the mathematical procedure suggested by Mould(11) and R. D. Southwick will be applid to the experimental data of the fatigue of glass, which has been reported Watanabe.(26)(3)

By taking the ratio of σ_{W}'/σ_{N}' at various values of the load duration, the "reduced fatigue curves" of glasses of different chemical compositions were obtained. Figs. 7, 8, 9 and 10 illustrate such reduced fatigue curves.

What is meant by the above mathematical procedure is the normalization of the values of the vertical axis (the strength tested in water) so as to indicate the degree of fatigue

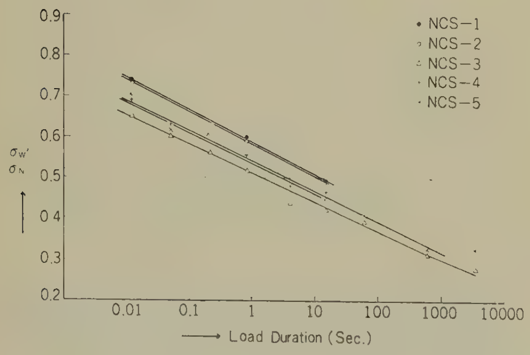


Fig. 7—Reduced fatigue curves for soda-lime glasses. (Part I.)

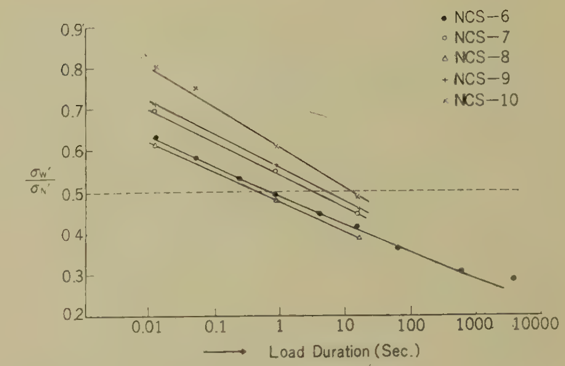


Fig. 8—Reduced fatigue curves for sodal-ime glasses. (Part II.)

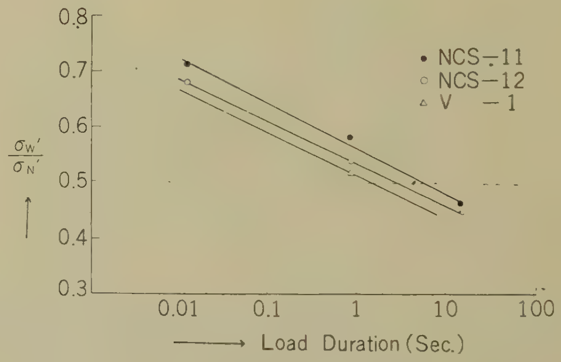


Fig. 9—Reduced fatigue curves of various glasses.

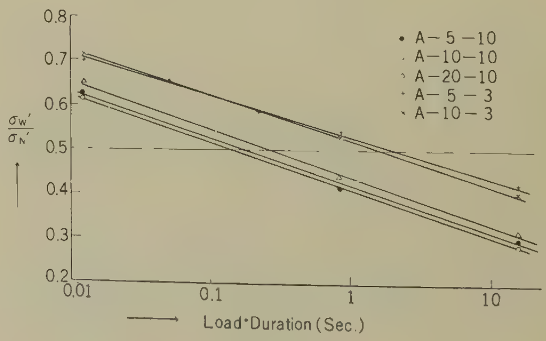


Fig. 10—Reduced fatigue curves for soda-lime glasses containing Al_2O_3.

of glasses of different chemical compositions.

In Figs. 7, 8, 9 and 10, the series of NCS and A glasses represent glasses of different chemical compositions in the Na_2O-CaO-SiO_2 ternary system and in the Na_2O-CaO-SiO_2-Al_2O_3 quarternary system.

Now, $t_{0.5}$ is defined as an indicator of the degree of fatigue, given as the characteristic load duration of each kind of glass when the value of σ_W'/σ_N' becomes 0.5. In other words, $t_{0.5}$ may well be regarded as a reaction constant of the physicochemical reaction between water and various glasses. Therefore, in (B) of Fig. 6, the time needed for the depth of the crack to increase from C_0' to C_{II} varies according to the different chemical compositions of the glasses.

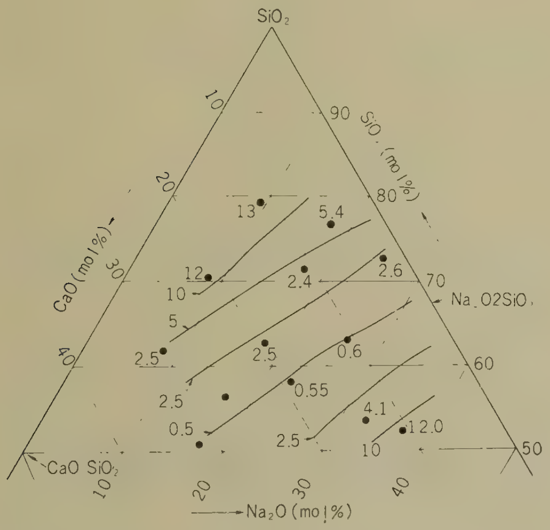


Fig. 11—Relation of $t_{0.5}$ (sec.) to various chemical compositions. An anomaly lies along the line between $Na_2O \cdot 2SiO_2$ and $CaO \cdot SiO_2$.

Fig. 11 shows some values of $t_{0.5}$, estimated from the reduced fatigue curves, for the NCS glasses. No additive relationship between $t_{0.5}$ and the composition of the oxide components such as Na_2O, CaO and SiO_2 is seen in this figure. A continuous deep valley of value $t_{0.5}$ lies along the *M* line which connects [Na_2O. $2SiO_2$] with [CaO. SiO_2]. This fact suggests that the mechanism of fatigue in the physicochemical reaction is not governed only by the simple oxide component, such as Na_2O.

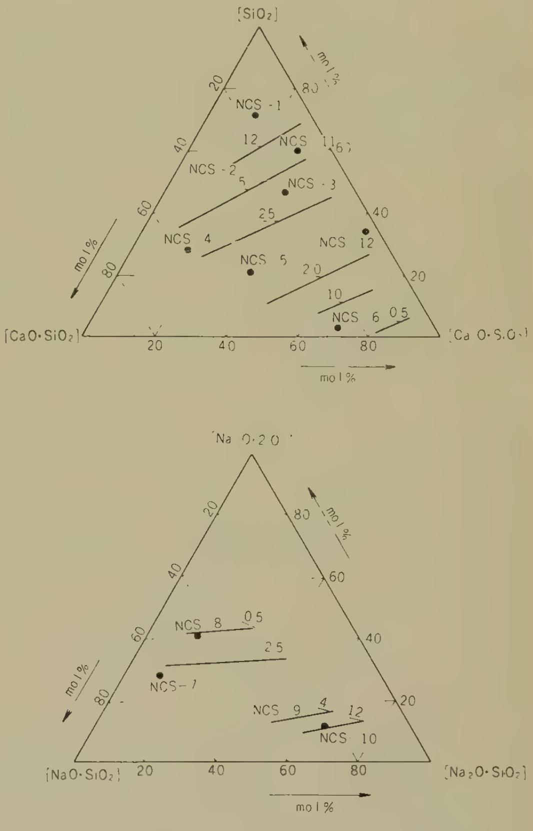


Fig. 12—Relation of $t_{0.5}$ (sec.) to various chemical compositions, based on the glassy microphase theory. No anomaly is seen.

Fig. 12 is a graphical presentation of $t_{0.5}$ according to the glassy microphase theory. Sufficient additivity is shown in these diagrams.

This fact leads us to a new idea on the mechanism of fatigue··the idea that $t_{0.5}$ is a reaction constant of the physicochemical reaction between water and the glassy microphase.

The idea is illustrated schematically by (C) and (D) of Fig. 6. In these figures, C_0' stands for the depth of the artificial abrasion, C_2 for the depth of the latent flaw mentioned in (C) and (D) of Fig. 3, and $C_{3\cdot1}$ and $C_{3\cdot2}$ represent the depths of other latent flaws caused by the physicochemical reaction between water and the glassy microphases under stress. Both $C_{3\cdot1}$ and $C_{3\cdot2}$ may well be considered to have been caused by H^+ ions and so forth in water which have penetrated into the glass and broken off the chemical bonds such as Si-O in the glassy microphases. Therefore it is considered that the flaw depths $C_{3\cdot1}$ and $C_{3\cdot2}$ cause the fatigue of glasses.

Now, another normalization for the horizontal axis of Figs. 7, 8, 9 and 10, was carried out by using $t/t_{0.5}$ instead of the simple load duration t.

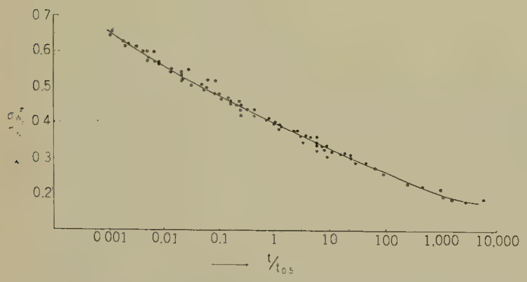


Fig. 13—A universal fatigue curve induced from the behavior of specimens of various chemical compositions with the same abrasion and aging conditions.

Fig. 13 shows the result of the normalization of the reduced fatigue curves. The figure clearly shows that many black dots represent the normalized values, and they form themselves into a curve. This curve coincides with the "universal fatigue curve" found by Mould(11) and Southwick. They induced it from their experiments on glass specimens of definite chemical composition with varied abrasion depths, whereas the authors induced it from the specimens of various chemical compositions with definite abrasion depth. The physical meaning of $t_{0.5}$ for both cases is different. The quantitative consideration of the universal fatigue curve in the latter case will be discussed in the following section.

The important point that should be emphasized here is that there is no other factor except σ_W'/σ_N' and $t_{0.5}$ in the universal fatigue curve which governs the mechanism of fatigue.

2.3.2. Considerations on the Universal Fatigue Curve

In order to make a mathematical formula for the universal fatigue curve, attention should be paid to the following terms.

(a) Since the normalized axis in Fig, 13 represents σ_W'/σ_N', the ratio of the strength in the presence of fatigue to the strength in the absence of fatigue should be conveniently taken at the left-hand side of the formula. The term σ_N' which stands for the strength (no fatigue) is given as follows, according to the Griffith formula, similar to Eq. (3):

$$\sigma_N' = k\sqrt{\frac{E\cdot\alpha}{C_0'+C_2}} \tag{6}$$

On the other hand, the strength σ_W' (fatigue) is given by the modified Griffith formula as follows:

$$\sigma_W' = k\sqrt{\frac{E\cdot\alpha}{C(t)}} \tag{7}$$

where $C(t)$ denotes that the depth of the flaw is a function of the load duration t.

From Eq. (6) and Eq. (7), σ_W'/σ_N' can easily be obtained.

(b) Vertical axis in Fig. 13 not represents the load duration t but $t/t_{0.5}$, where $t_{0.5}$ stands for the reaction constant characteristic to each kind of glass.

According to Moriya,(26) the deformation of a glassy microphase by an external force is due to the breakage of the chemical bonds between atoms in the microphase, and the deformation velocity of the glassy microphase under stress is proportional to the number of the bonds which have not yet broken at that time. Therefore, the velocity can be described as follows:

$$\frac{d\nu}{dt}=k_\nu(J-\nu) \qquad (8)$$

where $\frac{d\nu}{dt}$ = deformation velocity of the glassy microphase;

J = number of the chemical bonds to be broken by stress;

ν = number of the chemical bonds already broken at the time t;

k_ν = a constant which depends upon the shape, size and composition of the glassy microphases at various temperatures.

Watanabe(26) assumed that the same kind of formula might be applied to the reaction between water and the glassy microphases under stress; that the velocity of breaking the chemical bonds in the glassy microphases would be expressed by the number of the unbroken bonds at that time.

$$\frac{dn}{dt}=\alpha(N-n) \qquad (9)$$

where α = a constant which depends upon the shape, size and composition of the glassy microphases;

N = number of chemical bonds to be broken by the reaction;

n = number of the chemical bonds already broken at time t.

On the other hand, the breakage of chemical bonds by the reaction under stress produces a deep latent flaw. Provided that the number of the broken chemical bonds is proportional to C_3 which stands for the depth of this flaw,

$$\left.\begin{aligned} &C_3 \propto n \\ \therefore\ &C_3=p\cdot n \end{aligned}\right\} \qquad (10)$$

where p is a constant.

Substituting Eq. (10) into Eq. (9), we obtain

$$\frac{dC_3}{dt}=\beta'(H'-C_3) \qquad (11)$$

where β' and H' are constants.

Integrating Eq. (11), we obtain

$$C_3=H'(1-e^{-\beta' t}) \qquad (12)$$

In the actual breakage, the phenomenon is assumed to superposition of Eq. (12) as follows:

$$C_3=\sum_{j=1}^{i} H_j'(1-e^{-\beta_j' t}) \qquad (13)$$

Considering the value $t_{0.5}$ which is dependent on the difference of the chemical compositions of the glassy microphases, and using the corrected value β_j instead of the constant β_j', we obtain

$$C_3=\sum_{j=1}^{i} H_j'\{1-e^{-\beta_j(t/t_{0.5})}\} \qquad (14)$$

When $j=2$,

$$C_3=H_1\{1-e^{-\beta_1(t/t_{0.5})}\}+H_2\{1-e^{-\beta_2(t/t_0{}'_5)}\} \qquad (15)$$

Eq. (15) gives the depth of the latent flaw made by the reaction as a function of time.

Since the specimens used in the experiment by Watanabe(3) were initially abraded, the total depth of the flaw is

$$\begin{aligned} C_{\text{total}}&=C_0'+C_2+C_3 \\ &=C_4+C_3 \qquad (16) \end{aligned}$$

where $C_4=C_0'+C_2$.

By combining Eq. (16) with Eq. (15), we can get the time dependence of all flaws $C(t)$,

$$C(t)=C_4+H_1\{1-e^{-\beta_1(t/t_{0.5})}\}+H_2\{1-e^{-\beta_2(t/t_{0.5})}\} \qquad (17)$$

Substituting Eq. (17) into Eq. (7), and taking the ratio of Eq. (6) to Eq. (7), we obtain

$$\begin{aligned} \frac{\sigma_W'}{\sigma_N'}&=\left\{\frac{C_0'+C_2}{C(t)}\right\}^{\frac{1}{2}}=\left\{\frac{C_4}{C(t)}\right\}^{\frac{1}{2}} \\ &=\frac{C_4}{C_4+H_1\{1-e^{-\beta_1(t/t_{0.5})}\}+H_2\{1-e^{-\beta_2(t/t_{0.5})}\}} \qquad (18) \end{aligned}$$

Eq. (18) is an example of the final formula of the universal fatigue curve.

The values of the constants in Eq. (18) can be evaluated by the experimental data indicated in Fig. 13 as follows:

$$\left.\begin{array}{l} C_4=1.5\times10^{-4}(cm^{-1}) \\ H_1=3\times10^{-4}(cm^{-1}) \\ H_2=12\times10^{-4}(cm^{-1}) \\ \beta_1=107 \\ \beta_2=0.08 \end{array}\right\}$$

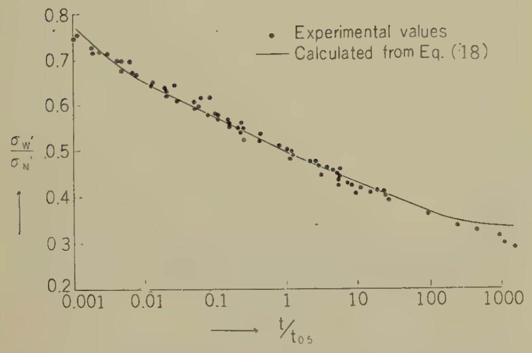


Fig. 14—Experimental points and a theoretical curve calculated from Eq. (18).

Substituting the above values into Eq. (18), we obtain the solid curve in Fig. 14, which shows good agreement with the experimental results.

In short, the authors wish to emphasize that the mechanism of fatigue is proved to be closely related to the glassy microphases, both experimentally and theoretically.

2.4. Miscellaneous Considerations on the Strength of Glass

(a) Load duration and the mirror area.

If the mechanism of fatigue can be represented as in Fig. 6, the relation between the mirror area on the surface of the fracture and the load duration can be easily explained. When fracture occurs after short load duration, the mirror area, the diameter of which is $(C_{\text{I}}-C_0')$ in (A) of Fig. 6, is smaller than that for a longer load duration. The comparison of (A) with (B), or (C) with (D) in Fig. 6 makes it clear that $(C_{\text{II}}-C_0')>(C_{\text{I}}-C_0')$. This means that a shorter load duration causes a smaller mirror area and greater strength. This consideration agrees well with a past(17) paper.

So, Watanabe predicts an appearance of the structure of the glassy microphase at the mirror part of the surface broken(22) with a moderate stress, because of the pattern as shown in cases of (C) and (D) of Fig. 6.

(b) Strength of Glass Fiber

The dependence of the strength of glass fiber on its diameter can be explained in the same way as in (A) and (B) of Fig. 3.

Glass fibers of the specimens quenched rapidly during the time of their formation have smaller glassy microphases than slowly cooled fibers, as shown in (A) and (B) of Fig. 3. Therefore, rapidly quenched glass fibers(25) with smaller diameters are stronger than slowly cooled glass fibers with larger diameters.

The above assumption also agrees well with the fact(13) that heat-treated glass fiber is weaker than non-heat-treated glass fiber.

(c) The difference between the physico-chemical meaning of $t_{0.5}$ used in the present paper and that used by Mould(11) is a project to be considered in the future.

3. Viscosity and Visco-elasticity of Glass within the Transformation Range

Since H. R. Lillie(19) found the so-called phenomenon of "viscosity change with time at a constant temperature," many experimental and theoretical reports have been presented to interpret the phenomenon.

Watanabe also reported(1)(2) some experimental results of the "large deformation" to show the relation between the load and the so-called viscosity, using the fiber-elongation method. However, he did not refer to the experimental result of so-called "small deformation," which can easily be treated theoretically.

In this paper, therefore, the authors wish to present the experimental results of the "small deformation" and an interpretatation of its relation to the internal structure of glass.

3.1. Experimental Results

All experimental results were obtained at a constant temperature, using the fiber-elongation method.

3.1.1. Results under a Constant Load

A load was attached to glass specimens at the begining of the experiment, and the weight of this load was held constant throughout the experiment. The elongation of the specimen with the elapse of time was measured by a traveling microscope. The "psuedo viscosity" $[\eta]$, which will be newly defined below, was calculated through the following equation using the experimental results.

$$[\eta]=\frac{1}{3}\cdot\frac{m\cdot g}{\frac{1}{l}\cdot\frac{dl}{dt}\cdot\pi r^2} \tag{19}$$

where

l = length of the specimen
m = load
g = specific gravity
r = radius of the specimen
$\frac{dl}{dt}$ = rate of the elongation of the specimen.

Usually the so-called viscosity η has been calculated from the right-hand side of Eq. (19), but the authors of the present paper wish to call this calculated value the "pseudo-viscosity."

For instance, the curves of the "psuedo-viscosity" changing with time in Fig. 15 (B) can be obtained by calculation from Eq. (19), using the values of the experimental points indicated in Fig. 15 (A).

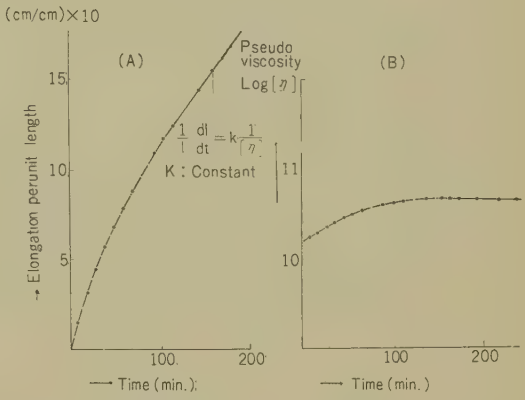


Fig. 15—The "pseudo-viscosity" change with time obtained from the experimental data of the elongation-method.

3.1.1.1. On Freshly Drawn Specimens

Figs. 16, 17, and 18 indicate the experimental results of the "pseudo-viscosity" of the W-1 (nonex type glass), S-1 (soda-lime

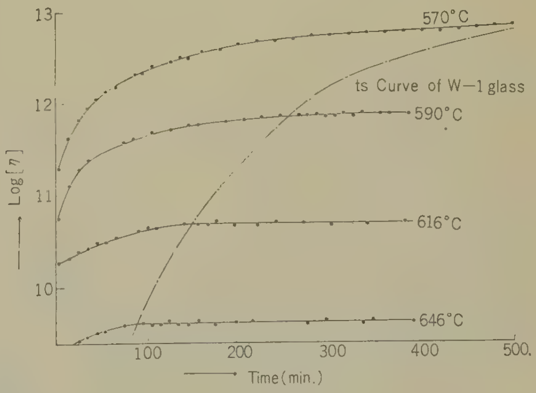


Fig. 16—The relations of the "pseudo-viscosity" to time for W-1 glass (nonex type glass). The t_s curve of W-1 glass.

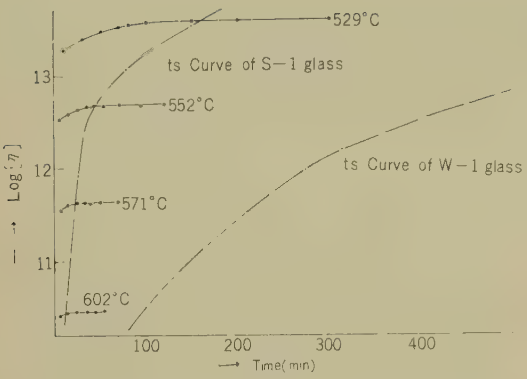


Fig. 17—The relation of the pseudo-viscosity to time for S-1 glass (sodalime glass). The t_s curve of S-1 glass.

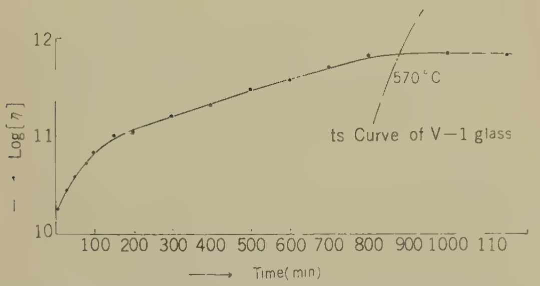


Fig. 18—The relation of the pseudo-viscosity to time for V-1 glass (Bolosilicate glass). The t_s curve of V-1 glass.

glass) and V–1 (borosilicate glass) glasses under a constant load.

These specimens are freshly drawn specimens, and have not been heat treated before the measurement.

It can clearly be seen from Fig. 16∼18 that the pseudo-viscosity curves of the several glasses are divided into two partsby t_s curves⁽¹⁾ the one dependent on time, and the other independent of time.

The time independent part of the values of the "pseudo-viscosity" can be immediately determined when the testing temperature is definitely chosen. For this reason the authors call this steady part of the pseudo-viscosity as the "steady viscosity". Since the steady viscosity has a linear relation to the reciprocal of the absolute temperature in a narrow temperature range, it can easily be understood to us.

The problem lies in the "pseudo-viscosity" (dependent on time), represented by the parts to the left of the t_s curves. The purpose of this section is to analyze this kind of the pseudo-viscosity.

Here a question may be raised whether the t_s curve of S–1 glasses and that of the V–1 glass are the same in character. The essential sameness in the meaning of both was proven in a previous paper.⁽¹⁾ The only difference lies in the retardation of the pseudo-viscosity according to different chemical compositions. Therefore the following analysis of the time dependency in the pseudo-viscosity to the left of the t_s curve is applicable to all kinds of glasses.

3.1.1.2. Specimens Heat Treated before the Test

The measurement of the pseudo-viscosity was carried out on the specimens which had been treated by heat beforehand under no stress, at the same constant temperature at which the freshly drawn specimens were tested. The obtained results are shown in Fig. 19.

In Fig. 19, curves (a), (b), (c), and (d) are the experimental results of the pre-heat-treat-

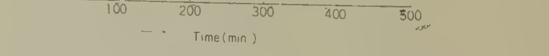


Fig. 19—Effects of pre-heating on the pseudo-viscosity of W-1 glass.

ed specimens, whereas curves (a)′ and (d)′ are those of the freshly drawn specimens, measured at the same constant temperature as indicated in Fig. 16. All curves of the pre-heat-treated specimens approach the value of the steady viscosity. This phenomenon of approach is independent of the conditions of the heat-treatment. But the time needed to approach the steady viscosity varies according to the conditions of the heat-treatment.

In general, heat treatment before the test is supposed to cause a chang in the internal structure of the glass. It follows that those parts of curves (a), (b) and (c) that are to the left of the t_s curve must correspond to the change of the internal structure of the W–1 glass at 616°C.

In other words, one of the causes of the change of the pseudo-viscosity with time before the t_s curve might well be the change of the internal structure of glass at a constant temperature. This is the important conclusion obtained from Fig. 19.

3.1.2. Additional Load Tests

A load in addition to the initial one was applied to the specimen at a selected time. The elongation rate of the specimen with time was measured, and its pseudo-viscosity was calculated from Eq. (19).

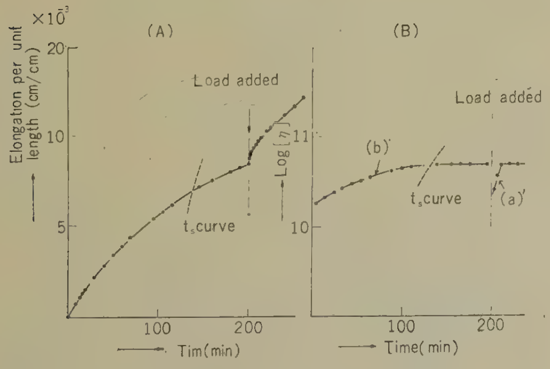


Fig. 22—Method of calculating the experimental results with additional load. Curve (a)′ is the visco-elasticity of glass based on the rheological model as shown in Figs. 20 and 21.

Fig. 22 (A) indicates the experimental data on the elongation of W–1 glass, with an additional load attached 200 minutes after the initial time. Fig. 22 (B) shows the logarithmic values of the pseudo-viscosity calculated from Fig. 22 (A) through Eq. (19).

As a previous paper⁽¹⁾ pointed out, the rheological "four element model" is a suitable for mechanical model interpreting the creep phenomenon of glass within the transformation range. When an additional load is applied to the model as illustrated in Fig. 20, the "instantaneous elasticity" is occured by E_1, the "visco-elasticity" by E_2 and η_2, and the "steady viscosity" by η_3.

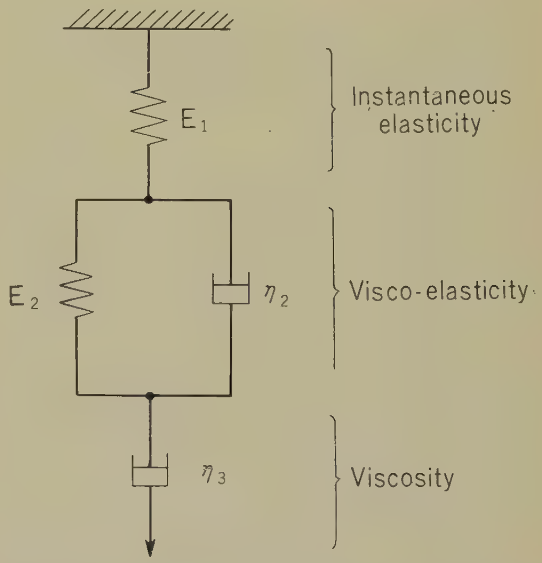


Fig. 20—The rheological model. (Four element model).

Fig. 21 is a schematic diagram of the actual deformation of the specimen under the rheological test. The curves drawn between the points O-(H), (H)-(I), and (I)-(J) represent the above three behaviors of glass.

An important point in Fig. 21 is that the time interval during which the behavior of the visco-elasticity is seen by the initial load, $(t_{\lambda\text{-}1})$, is equal to the time interval during which the behavior of visco-elasticity is con-

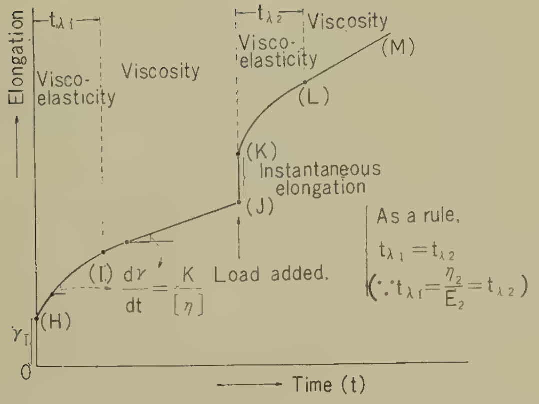


Fig. 21—Elongation vs. time (according to the rheological model).

tinued by the additional load, $(t_{\lambda\cdot1})$. This phenomenon is true as long as the structure of glass is unchanged (subsequently E_2 and η_2 is not the function to time at a constant temperature.)

A comparison of Fig. 22 (A) with Fig. 21 clarifies that the tendency of the elongation curve of the glass specimen is similar to the rheological elongation curve. However, in Fig. 22 (A), the time interval in which the visco-elasticity is seen by the initial load and that by the additional load are not the same. Namely, the former is about 130 minutes, as indicated by (b)′ of Fig. 22 (B), whereas the latter is only about 10 minutes, as shown by (a)′ of Fig. 22 (B).

The important question is which curve, (a)′ or (b)′ in Fig. 22 (B), represents the true visco-elasticity at this temperature. In the opinion of the authors, it is (a)′ of Fig. 22 (B) that denotes the true visco-elasticity of the glass at this temperature. The reason is that (a)′ appears after stabilization of the internal structure of glass, whereas (b)′ is affected by the still-persisting change of the internal structure of the glass, caused by the heat-treatment before the test, as was described above for Fig. 19.

3.1.2.1. Effects of the Load Added after the Time t_s

Fig. 23 and 24 show the true visco-elastic behavior of glasses caused by additional load applied after the time t_s, i.e. after the stabilization of the internal structure. These figures clarify the following points.

(a) The visco-elastic behavior appears for glasses of various chemical compositions at various temperatures, as shown by (a)′ and (b)′ in Figs. 23 and 24.

(b) Steady viscosity was observed after the visco-elastic behavior appeared by ad-

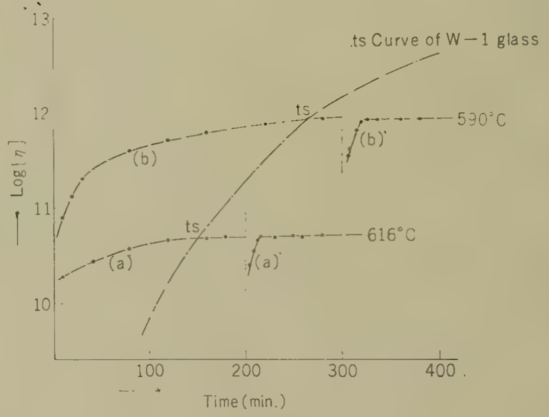


Fig. 23—True visco-elasticity of W-1 glass, which appears after the additional load is applied; on the right-hand side of the t_s curve.

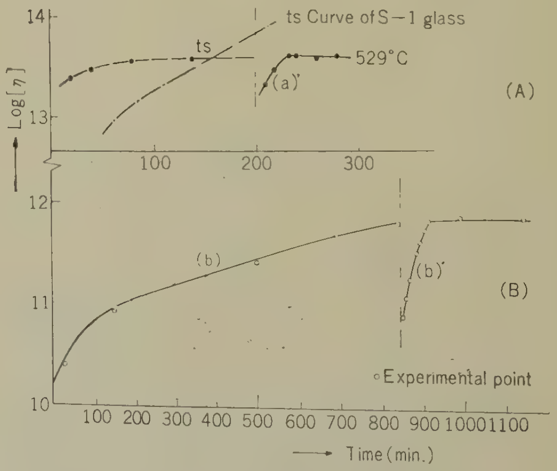


Fig 24—True visco-elasticity of V-1 glass several glasses.

ditional load. The values of this following steady viscosity is same# to the steady viscosity of the freshly drawn specimens.

(c) The time interval which the behavior of the visco-elasticity continues due to the additional load is much shorter than the time t_s.

3.1.2.2. Effects of the Load Added before Time t_s

As stated in 3.1.2, the change of the pseudo-viscosity with time before the time t_s is not the true visco-elasticity. The question may occur, however, whether the true visco-elasticity exists before the time t_s or not.

Fig. 25 is an answer to this question. The obtained results from Fig. 25 are as follows.

(a) The visco-elastic behavior indicated by (a)″, (b)″, and (d)″ does appear, due to the additional load applied before time t_s.

(b) After the appearance of the above visco-elasticity, the change of the pseudo-viscosity with time is still observed. This suggests the fact that the change of the internal structure of the glasses still continues toward a steady state.

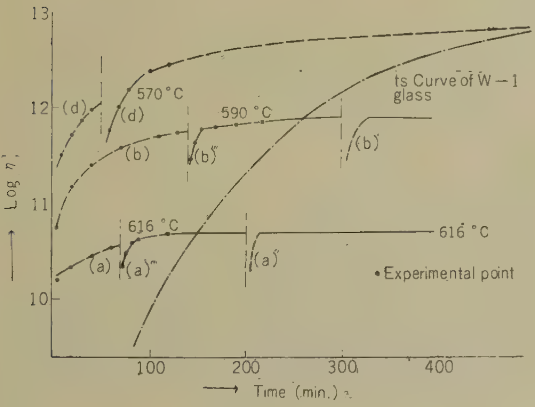


Fig. 25—Visco-elasticity of W-1 glass which appears after the additional load is applied; on the left-hand side of the t_s curve. Even in this region, the visco-elastic behavior, (a)″, (b)″, can be observed, and these are the same as (a)′, (b)′.

Summarizing all of these experimental results, the authors wish to emphasize that the change of the pseudo-viscosity with time before the time (t_s) must have been caused by the superposition of the time change of the internal structure of glass and the visco-elastic behavior. The reasons are that the pseudo-viscosity changed by the heat-treatment before the test, reflecting the change of the internal structure of glass (Fig. 19), and also that the true visco-elastic behavior due to the additional load applied before the time t_s was observed. (Fig. 25).

In Table 1, the pseudo-viscosity is classified according to various kinds of mechanisms. This was carried out by taking into account all the experimental results given in this section. Namely, the pseudo-viscosity is classified into several phenomena according to the kinds of load applied and of the time dependence. In the final stage, the pseudo-viscosity was divided into pure viscosity, pure visco-elasticity, and the superposition of the visco-elasticity and the time change due to the internal structure.

The Table 1 gives us a clear image of all the experimental results, and makes it more comprehensible that the so-called viscosity change with time reported in the past papers (1)(2) is in fact caused by the superposition of the two phenomena, as mentioned above.

3.2. A Consideration on the Time Change of Viscosity

3.2.1. The "Glassy Microphase" Theory Proposed by Moriya

In a paper by Moriya(8), the phenomenon of the superposition was predicted theoretically and experimentally. In that paper, the growth of the glassy microphase with time at a constant temperature was identified with the change of the internal structure of the glass, and statistical thermodynamics was used to introduce a mathematical formula for

---

# This holds true for the experiments of the so-called "small deformation," due to alight load. But it deviated a little for the experimental of the "large deformation," due to a heavier load. This difference was already pointed out in one of Watanabe's papers.(1)

Table 1

CLASSIFICATION OF "PSEUDO-VISCOSITY"

Experimental quantity	Load system	Change with time		Classified name	Assumed mechanism
		Before t_s.	After t_s.		
"Pseudo-viscosity"	Constant load	yes.		So-called viscosity change with time.	Superposition of viscosity and visco-elasticity while the internal structure is changing.
			No.	Steadyviscosity.	Viscosity after the internal structure was stabilized.
	Additional load	yes.		Visco-elasticity.	Visco-elasticity while the internal structure is changing.
			yes.	Visco-elasticity and viscosity.	Visco-elacticity; after the internal structure was stabillzed. Viscosity; when the internal structure was stabilized.

the time change of viscosity at constant temperature.

This formula was examined in the light of the experimental data presented by H. R. Lillie(18), and good agreement was obtained between the theoretical curve and the experimental data.

In this paper, it is again necessary to use this mathematical formula to explain the authors' idea. Therefore it will be described in an abbreviated manner.

Viscosity η may be defined as,

$$\eta=\frac{S}{\frac{d\gamma_1}{dt}+\frac{d\gamma_2}{dt}}, \tag{20}$$

where $\frac{d\gamma_1}{dt}=$ elongation rate due to the flow of the glassy microphase,

$\frac{d\gamma_2}{dt}=$ elongation rate due to the visco-elastic behavior of the glassy microphase,

$S=$ stress.

According to the Eyring's theory,(19) the first term in the denominator of Eq. (20) is

$$\frac{d\gamma_1}{dt}=KVS_1 \tag{21}$$

where $K=$ constant

$V=$ volume of a flow unit.

On the other hand, the volume of glassy microphases is supposed to change with time at a constant temperature as follows:

$$V(t)=V_T(1-e^{-\alpha t})+V_0e^{-\alpha t} \tag{22}$$

where $V_T=$ volume of glassy microphases at temperature T, and time $t\rightarrow\infty$,

$V_0=$volume of glassy microphases at temperature T, and time $t=0$,

$\alpha=$constant.

Provided that the initial volume of "frozen in glassy microphases" is V_τ, Eq. (22) is changed as follows:

$$V(t)=V_T-(V_T-V_\tau)e^{-\alpha t} \tag{23}$$

The temperature dependence of the growth of the glassy microphase is:

$$V_T=A''-Be^{-\Delta E/T} \tag{24}$$

where ΔE=the energy term according to Maxwell-Boltzman statistics, and A'' and B are constants.

Substituting Eq. (24) into Eq. (23), we can obtain a detailed formula for $V(t)$. Then by substituting the detailed formula of $V(t)$ into Eq. (21), a detailed formula for $\frac{d\gamma_1}{dt}$ can be obtained. The resulting formula shows the change of the flow of glassy microphases with time.

The Voigt-model from rheology is applied to the visco-elastic behavior for clarification of the flow, $\frac{d\gamma_2}{dt}$.

By substituting the detailed formulas of $\frac{d\gamma_1}{dt}$ and $\frac{d\gamma_2}{dt}$ into Eq. (20), we obtain the following final formula:

$$\eta=\frac{S}{KS_1\{V_T+Be^{-\Delta E/T}(1-e^{-\Delta E\cdot\frac{\tau-T}{\tau\cdot T}})e^{-\alpha_1 t}\}+J\cdot e^{-\alpha_2 t}} \tag{25}$$

where K, B, α_1, α_2 and J are constants, and τ=fictive temperature.

This equation shows the dependency of viscosity on time and temperature.

At a constant temperature, Eq. (25) can be written as:

$$\eta=\frac{1}{a\pm be^{-\alpha_1 t}+de^{-\alpha_2 t}} \tag{26}$$

where a, b, d, α_1 and α_2 are constants. It is worth mentioning that the first and second terms in the denominator were introduced from the volume change of glassy microphases with time, and the third term was introduced from the visco-elastic behavior of the glassy microphase.

The authors wish to apply this principle to the experimental results on the growth of the glassy microphase, performed by means of an electron microscope, as well as to those on the visco-elasticity of the same glass. These will be described in a later section, (3.2.2.).

When Moriya introduced the third term of the denominator in Eq. (25) from $\frac{d\gamma_2}{dt}$ in Eq. (20), another procedure was also available to obtain the same formula. This point is elaborated below.

Moriya(26) assumed that the angle of a chemical bond (or the rotational angle) in glassy microphases would change with applied stress, and that their deformation rate $\frac{d\theta}{dt}$ could be expressed as:

$$\frac{d\theta}{dt}=k_\theta(\Theta-\theta) \tag{27}$$

where Θ=total number of the bond angles (or rotational angles) that would be changed by an applied stress,

θ=number of the moved bond angles (or rotational angles) at time t,

k_θ=a constant depending upon the structure and composition of glassy microphases at various temperatures.

In the same manner, the deformation rate of glassy microphases can be expressed by Eq. (8).

Integrating Eq. (27) and Eq. (8) with the boundary conditions of $\theta=0$ and $\nu=0$ at $t=0$, we have,

$$\theta=\Theta(1-e^{-k_\theta}) \tag{28}$$

$$\nu=J(1-e^{-k_\nu}) \tag{29}$$

Both Eqs. (28) and (29) contribute to the elongation (or deformation) of the glassy microphase. The total elongation γ_m can be written as:

$$\gamma_m=\sum_m H_m(1-e^{-\alpha_m t}) \tag{30}$$

where H and α_m are constants.

Experimentally, only the mean value of Eq. (30) may be obtainable. The mean value of the elongation of glassy microphases in the visco-elastic effect, γ_2, is

$$\gamma_2=H(1-e^{-\alpha_2 t}) \qquad (31)$$

The derivative of Eq. (31) is

$$\frac{d\gamma_2}{dt}=H\alpha_2 e^{-\alpha_2 t}$$

$$=F\cdot e^{-\alpha_2 t} \qquad (32)$$

Obviously Eq. (32) is essentially the same as the third term of the denominator in Eq. (26), in spite of the difference of the procedures used to obtain them.

3.2.2. Application of Moriya's Theory

As stated before, the first and second terms of the denominator in Eq. (26) were introduced from the change of the glassy microphase with time. Therefore the experimental curve of Fig. 2 (a) can be used to determine the constants in these terms for V-1 glass.

The third term of the denominator in Eq. (26) was introduced from the visco-elastic behavior of glass. Therefore, experimental curve (b)′ in Fig. 24 can be used to determine the constants in this term for the same glass at the same temperature.

Substituting these constants into Eq. (26), we obtain a theoretically calculated curve, which is useful to explain the viscosity change with time.

The actual process of the calculation is as follows:

(a) By assuming the existence of a proportionality relation between the volume of glassy microphases and the cube of the "width" in curve (a) of Fig. 2, the constant α_1 of the experimental formula $a+be^{-\alpha_1 t}$ can be determined. Since $T-\tau<0$ in this case, the sign of b is plus.

(b) The constant α_2 in $de^{-\alpha_2 t}$ is obtained from the curve (b)′ in Fig. 24.

(c) Substituting these constants into Eq. (26), we obtain the theoretically calculated curve of the viscosity change with time.

Fig. 26 illustrates this calculated curve, which agrees well with the experimental points.

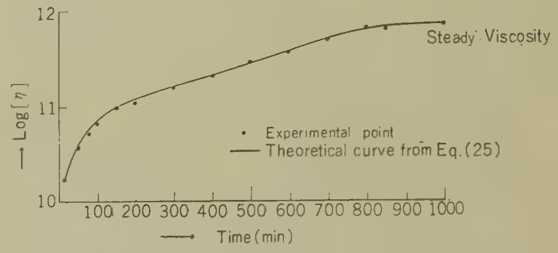


Fig. 26—Experimental points and the calculated curve.

A similar calculation procedure is applied to (a)′ and (b)′ of Fig. 23 (W-1 glass) and to (a)′ of Fig. 24 (S-1# glass). The constants obtained are listed in Table 2. And the theoretical curves mentioned above agree well with the experimental data, as shown in Fig. 27.

In short, the authors with to emphasize that Eq. (26), based on the glassy microphase theory, is a highly valid equation to interpret the viscosity change with time at various temperatures and for various kinds of glass.

Summary and Conclusions

(i) On the Strength of Fatigueless Glass

The authors regarded the part of the weaker chemical bond existing in the bondary layer between a glassy microphase and matrix

---

# Another evaluation was adapted from the data of H.R. Lillie at 477°C, 1,380 min. Using the same calculation procedure as above, the authors obtained 0.007 as the value of α_2 for soda lime glass for the additional load (1,380 min.). The value agrees well with $\alpha_2=0.0065$ given by Moriya. The important point is that the both values were obtained through different methods but agreed well with each other. Therefore, α_2 is a factor which should be introduced from the visco-elastic behavior of glass.

Table 2

THE VALUES OF CONSTANTS IN EQ. (26)

Temperature (°C)	a	b	d	α_1	α_2
V-1 glass					
570	0.14×10^{-10}	1.6×10^{-10}	3.26×10^{-10}	0.0027	0.038
W-1 glass					
616	0.2×10^{-10}	0.45×10^{-10}	0.35×10^{-10}	0.0081	0.089
590	0.12×10^{-11}	0.33×10^{-11}	1.75×10^{-11}	0.0046	0.076
570	0.12×10^{-12}	0.27×10^{-12}	5.6×10^{-12}	0.0013	0.046
S-1 glass					
529	2.53×10^{-14}	3.2×10^{-14}	0.9×10^{-14}	0.0006	0.022
(Reference) According to T. Moriya, the constants for the glass on which H. R. Lillie experimented were as follows,					
477°C	0.36×10^{-15}	3.54×10^{-15}	0.56×10^{-15}	0.000096	0.0065

Watanabe obtained $\alpha_2=0.007$ by calculation, appling Eq. (26) to his experimental results on visco-elasticity, 477°C load added, 1,381 mins.~2,000 mins., after. Both values agree with each other very well.

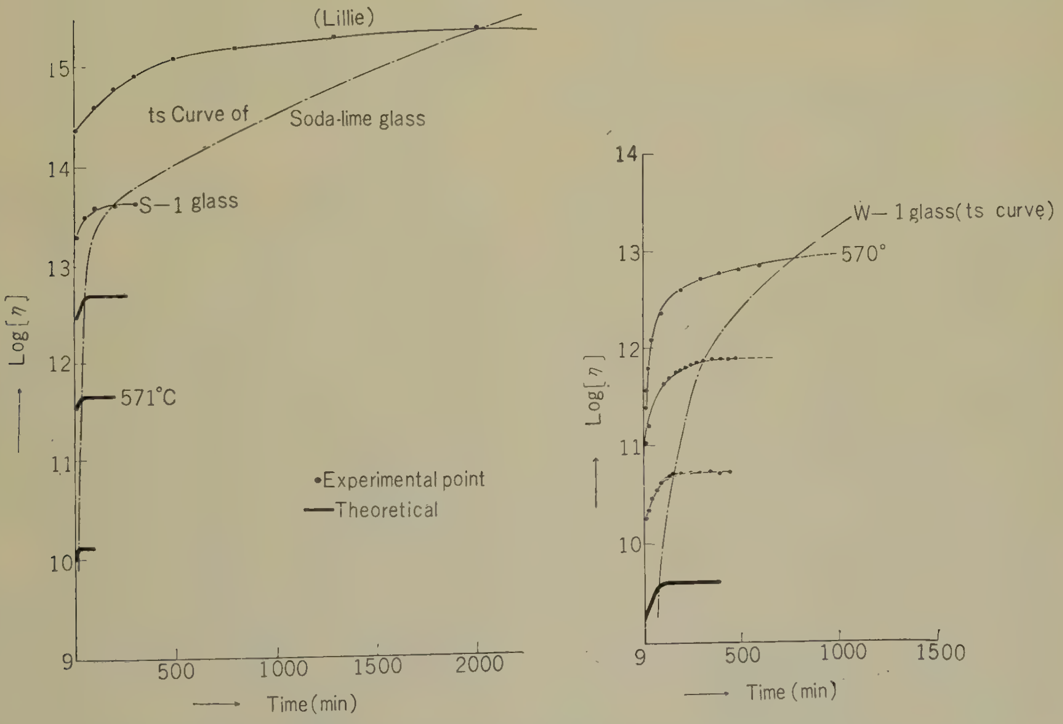


Fig. 27—Experimental data and calculated curves for W-1 and S-1 glasses.

part (completely amorphous) as a latent flaw which may safely be considered the Griffith's flaw. From this consideration, Griffith's formula may be only applied to the strength of glasses without fatigue.

For artificially abraded glass, the authors assumed that the size of the glassy microphase at the tip of the abrasion governs the strength of the glass.

(ii) On the Strength of Fatigued Glass

The authors assumed that the fatigue phenomenon was caused by a physicochemical reaction between water and glassy microphases under stress; that the latent flaw would grow during the reaction, and cause fatigue.

The time of the reaction between water and glasses of various chemical compositions, were defined as $t_{0.5}$. Further, the authors obtained a universal fatigue curve, which was given by Mould and Southwick.

And a mathematical method to introduce the universal fatigue curve was theoretically shown.

(iii) On the Flow Properties of Glasses within the Transformation Range

From the fiber-elongation method, the "pseudo-viscosity" was defined as an experimental quantity. Then it was concluded that the "pseudo-viscosity" can be classified into the phenomena of viscosity and of visco-elasticity, according to the method of appling stress and the conditions of the heat treatment of the glasses before the test.

The behavior of the pseudo-viscosity which appears for a short time by appling additional load was concluded to be the true visco-elastic behavior of glass.

An important conclusion was obtained from these experiments. Namely, the behavior of viscosity change with time under constant load at a constant temperature for freshly drawn specimens, has two superposed part, one is pure visco-elastic behavior and the other due to the change of the internal structure of glass.

On the basis of this conclusion, both the electron microscopic and the visco-elastic experimental results were applied to the equations deduced from the glassy microphase theory. As a result, a theoretical curve of the "viscosity change with time" was obtained for a borosilicate glass.

This mathematical procedure was applied also to other kinds of borosilicate glasses and to soda lime glasses at various temperatures. These theoretical curves coincided with the experimental data very well.

In short, the glassy microphase theory is considered a highly valid interpretation of the internal structure of glasses.

Acknowledgement

The authors would like to give thanks to Dr. Kojima, Director of the Electrical Communication Laboratory, and to Dr. Hirabayashi, Chief of the Ferro-electric Section, for permission to publish this paper.

The authors are also indebted to Assistant Professor T. Sakaino of The Tokyo Institutes of Technology for his helpful discussions, and to Mr. T. Aiba who carried out a part of the calculation.

Appendix

It is necessary to show the experimental evidence for the existence of the glassy microphase in every kind of glass. Two methods are available for this purpose: one is the X-ray small angle scattering method and the other is the electron microscope method. However, only a few reports(5), (20) have been presented on these experiments. In the future, the experimental results reported by J. W. Michener(21) and A. F. Prebus, J. Warshow(22) and Hoffman(23), as well as some other experimental results(24) would be used to endorse the glassy microphase theory.

Fig. 28 is only one example of the schematic representation of the change of the glassy microphase, and Fig. 1 is another. These patterns will vary with the chemical compositions of glassy-microphase and with temperature. The method of expressing the varied patterns is a project to be considered in the future.

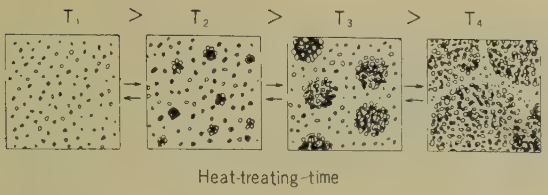


Fig. 28—Another schematic representation of the change of the internal structure of glass.

References

(1) M. Watanabe and R. Koyama; *J. Soc. Glasstech.*, **41**, 137, (1957).

(2) M. Watanabe, T. Chida and K. Takizawa; *J. Soc. Glasstech.*, **41**, 295 (1957).

(3) M. Watanabe, R. V. Caporali and R. E. Mould; This report will be published in the Physics and Chemistry of Glass. (Feb. 1961).

(4) M. Watanabe; *Electrical Communication Laboratory Reports*, **720**, (1959). (In Japanese.)

(5) M. Watanabe, H. Noake and T. Aiba; *J.Am. Ceram. Soc.*, **42**, 593 (1959).

(6) M. Watanabe, H. Noake and T. Aiba; *Electrical Communication Laboratory Reports*, **1,021**, (1957). (In Japanese.)

(7) T. Moriya; *J. Japanese Ceram. Association*, **55**, 60, 87, 134 (1947) and **56**, 4 (1948).

(8) T. Moriya; *Bull. of Tokyo Institutes of Tech.*, B. **2**, 47 (1955).

(9) A. A. Griffith; *Phil. Trans. Roy. Soc.* A., **221**, 163 (1920).

(10) F. W. Preston; *J. Appl. Phys.*, **13**, 623 (1942).

(11) R. E. Mould and R. D. Southwick; *J. Am. Cer. Soc.*, **42**, 582 (1959).

(12) T. C. Baker and F. W. Preston; *J. Appl. Physics*, **17**, 179 (1946).

(13) S. Sakuka; *J. Japanese Ceram. Association*, **65**, 190 (1957).

(14) S. Spinner; *J. Am. Cer. Soc.*, **39**, 113 (1956).

(15) C. W. Parmee; *Cer. Abst.*, **18**, (11), 295 (1939).

(16) R. E. Mould; *J. Am. Cer. Soc.*, **43**, 160 (1960).

(17) W. C. Levengood; *J. Appl. Physics.*, **29**, 820 (1958).

(18) H. R. Lillie; *J. Am. Cer. Soc.*, **16**, 619 (1933).

(19) H. Eyring; "*The Theory of Rate Process*," New York: McGraw-hill Co., (1941).

(20) K. Kuehne und W. Skatulla; *Silikat-technik*, **10**, No. 3, 105 (1959).

(21) J. W. Michener and A. F. Prebus; *Ind. Eng. Chem.*, **46**, 147 (1954).

(22) I. Warshaw; *J. Am. Cer. Soc.*, **43**, 4 (1960).

(23) L. C. Hoffman and W. O. Statton; *J. Appl. Physics*, **31**, 404 (1960).

(24) M. Watanabe and H. Noake; to be published.

(25) W. H. Otto, *J. Am. Cer. Soc.*, **38**, 122 (1955).

(26) M. Watanabe, T. Moriya; *J. Japanese Ceram. Association*, **68**, 11, 273, (1960), **69**, 1, (1961).

* * * *

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 1-2, JANUARY-FEBRUARY, 1961

U.D.C. 621. 318. 5. 001. 2

Effects of the Force Factor and the Eddy Current on the Operate and Release time of the Ordinary Relay[*]

Yasuo TOMITA[†]

The eddy current effect can not be neglected in the motional analysis of relays, especially, in the analysis of open-release motion. The author derived motional equations taking the effect of eddy current into consideration, and the theoretical values of operate and release time were compared with the experimental ones. The constants in the motional equations that represent the effect of eddy current can be obtained by force-factor measurement. Although the force-factor measurement of ordinary relays is not common, this is one of the important items necessary for the motional analysis of these devices. This paper describes the results obtained.

1. Introduction

The eddy current effect is very important, in comparison with hysterisis and skin effects in the analysis of relays; especially in the analysis of the release motion. The fundamental motional equations considering the effect of eddy current have been derived from the Lagrange-Maxwell equation, and the experimental values of operate and release times have been compared with the theoretical values.

Though force factor is not ordinarily considered in the motional analysis of the non-polarized devices, it is very important and useful for analysis of the non-polarized devices as well as for that of the polar devices. Some data of force factor will be shown in this paper.

2. The Fundamental Equations

Fig. 1 illustrates a model of the structure of a magnetic circuit in which the eddy cur-

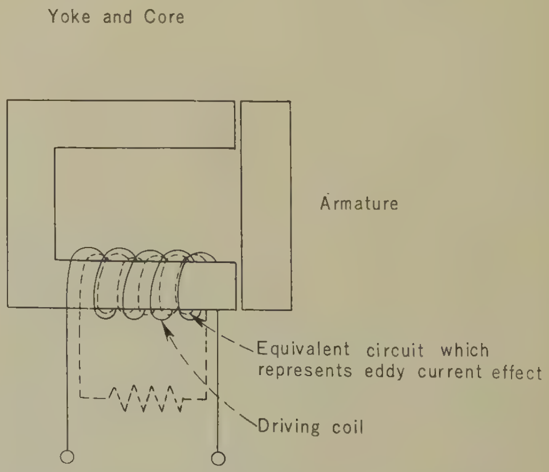


Fig. 1—Model of the structure of the ordinary relay in which the eddy-current effect is included.

---

[*] MS in Japanese received by the Electrical Communication Laboratory, on 19, July, 1960. Originally published in *The Journal of the Institute of Electrical Communication Engineers of Japan*, Vol. 44, No. 1, pp. 21–19, 1961.

[†] Station Apparatus Research Section.

rent effect is considered. The fundamental equations can be obtained from the Lagrange-Maxwell equation, all of whose terms including the kinetic, electro-magnetic, and potential energies are represented in the structure of Fig. 1.

Hence:

$$\left.\begin{aligned} E&=L_1\frac{dI_1}{dt}+L_{12}\frac{dI_2}{dt}+\Phi_1\frac{dx}{dt}+R_1I_1\\ O&=L_2\frac{dI_2}{dt}+L_{12}\frac{dI_1}{dt}+\Phi_2\frac{dx}{dt}+R_2I_2\\ F&=m\ddot{x}+r\dot{x}+sx-\frac{1}{2}\Phi_1I_1-\frac{1}{2}\Phi_2I_2 \end{aligned}\right\} \quad (1)$$

where, Φ_1 and Φ_2 are the so-called force factors.

$$\left.\begin{aligned} \Phi_1&=\frac{\partial L_1}{\partial x}I_1+\frac{\partial L_{12}}{\partial x}I_2\\ \Phi_2&=\frac{\partial L_2}{\partial x}I_2+\frac{\partial L_{12}}{\partial x}I_1 \end{aligned}\right\} \quad (2)$$

where

F=v.m.f., E=e.m.f., x=displacement, r=loss factor, m=mass, s=stiffness, L_1, L_2, L_{12}=equivalent self and mutual inductance of the driving coil and eddy current circuit, I_1=driving current, I_2=equivalent eddy current, R_1=resistance of driving coil, R_2=equivalent resistance of eddy current circuit.

And attractive force F is expressed as follows:

$$F=\frac{A(4\pi N_1I_1+4\pi N_2I_2)^2}{8\pi(AR_0+x_g)^2(1-u)^2} \quad (3)$$

where

$$u=\frac{x}{AR_0+x_g}. \quad (4)$$

Inductance and force factor are as follows:

$$\left.\begin{aligned} L_1&=\frac{4\pi N_1^2A}{(AR_0+x_g)(1-u)}\\ L_2&=\frac{4\pi N_2^2A}{(AR_0+x_g)(1-u)}\\ L_{12}&=\frac{4\pi N_1N_2A}{(AR_0+x_g)(1-u)} \end{aligned}\right\} \quad (5)$$

$$\left.\begin{aligned} \Phi_1&=\frac{4\pi N_1^2AI_1+4\pi N_1N_2AI_2}{(AR_0+x_g)^2(1-u)^2}\\ \Phi_2&=\frac{4\pi N_2^2AI_2+4\pi N_1N_2AI_1}{(AR_0+x_g)^2(1-u)^2} \end{aligned}\right\} \quad (6)$$

where

N_1, N_2: number of turns corresponding to L_1, L_2,
A : effective pole-face area,
R_0 : closed gap reluctance,
x_g : main gap length.

Equations (1) can be modified into simpler and non-dimensional ones by means of the the substitution of the following expressions:

$$\left.\begin{aligned} I_1&=\frac{Ei_1}{(L_1)_0\omega_0}, & I_2&=\frac{Ei_2}{(L_2)_0\omega_0}\\ \beta_1&=\frac{R_1}{(L_1)_0\omega_0}, & \beta_2&=\frac{R_2}{(L_2)_0\omega_0}\\ \sigma_1&=\frac{(L_{12})_0}{(L_2)_0}=\sqrt{\frac{(L_1)_0}{(L_2)_0}}=\frac{1}{\sigma_2} \end{aligned}\right\} \quad (7)$$

where the parentheses $(\ )_0$ indicate the initial value.

$$\left.\begin{aligned} \frac{A(4\pi N_1I_1)^2}{s8\pi(AR_0+x_g)^3}&=\Delta\beta_1^2i_1^2\\ \frac{A(32\pi^2N_1N_2I_1I_2)}{s8\pi(AR_0+x_g)^3}&=\Delta'\beta_1\beta_2i_1i_2\\ \frac{A(4\pi N_2I_2)^2}{s8\pi(AR_0+x_g)^3}&=\Delta''\beta_2^2i_2^2. \end{aligned}\right\} \quad (8)$$

Then the operating equations can be rewritten as follows:

1) The waiting range is:

$$\left.\begin{aligned}&\frac{di_1}{d\theta}+\sigma_1\frac{di_2}{d\theta}+\beta_1 i_1-1=0\\&\frac{di_2}{d\theta}+\frac{1}{\sigma_1}\cdot\frac{di_1}{d\theta}+\beta_2 i_2=0\\&\Delta\beta_1^2 i_1^2+\Delta'\beta_1\beta_2 i_1 i_2+\Delta''\beta_2^2 i_2^2=u_0\end{aligned}\right\}\tag{9}$$

$u_0=u_{x=x_0}$, $x_0=\frac{F_0}{s}$, $F_0=$back tension.

2) The motional range is:

$$\frac{1}{1-u}\cdot\frac{di_1}{d\theta}+\frac{\sigma_1}{1-u}\cdot\frac{di_2}{d\theta}+\frac{i_1}{(1-u)^2}\cdot\frac{du}{d\theta}+\frac{i_2}{(1-u)^2}\sigma_1\frac{du}{d\theta}+\beta_1 i_1-1=0\tag{10}$$

$$\frac{1}{1-u}\cdot\frac{di_2}{d\theta}+\frac{1}{(1-u)\sigma_1}\cdot\frac{di_1}{d\theta}+\frac{i_2}{(1-u)^2}\cdot\frac{du}{d\theta}+\frac{i_1}{(1-u)^2\sigma_1}\cdot\frac{du}{d\theta}+\beta_2 i_2=0\tag{11}$$

$$\frac{d^2u}{d\theta^2}+2\delta_0\frac{du}{d\theta}+u+u_0=\frac{1}{(1-u)^2}(\Delta\beta_1^2 i_1^2+\Delta'\beta_1\beta_2 i_1 i_2+\Delta''\beta_2^2 i_2^2).\tag{12}$$

Next, the release equations can be expressed as follows:

1) The waiting range of open release is:

$$\left.\begin{aligned}&\frac{1}{1-u_g}\cdot\frac{di_2}{d\theta}+\beta_2 i_2=0\\&\Delta''\beta_2^2 i_2^2=(u_g+u_0)(1-u_g)^2.\end{aligned}\right\}\tag{13}$$

2) The motional range of open release is:

$$\frac{1}{1-u}\cdot\frac{di_2}{d\theta}+\frac{i_2}{(1-u)^2}\cdot\frac{du}{d\theta}+\beta_2 i_2=0\tag{14}$$

$$\frac{d^2u}{d\theta_2}+2\delta_0\frac{du}{d\theta}+u+u_0=\frac{\Delta''\beta_2^2 i_2^2}{(1-u)^2}.\tag{15}$$

3) The waiting range of short circuit release is:

$$\left.\begin{aligned}&\frac{1}{1-u_g}\cdot\frac{di_1}{d\theta}+\frac{\sigma_1}{1-u_g}\cdot\frac{di_2}{d\theta}+\beta_1 i_1=0\\&\frac{1}{1-u_g}\cdot\frac{di_2}{d\theta}+\frac{1}{(1-u_g)\sigma_1}\cdot\frac{di_1}{d\theta}+\beta_2 i_2=0\\&\Delta\beta_1^2 i_1^2+\Delta'\beta_1\beta_2 i_1 i_2+\Delta''\beta_2^2 i_2^2\\&\quad=(n_g+u_0)(1-u_g)^2.\end{aligned}\right\}\tag{16}$$

4) The motional range of short circuit release is:

$$\frac{1}{1-u}\cdot\frac{di_1}{d\theta}+\frac{\sigma_1}{1-u}\cdot\frac{di_2}{d\theta}+\frac{i_1}{(1-u)^2}\cdot\frac{du}{d\theta}+\frac{i_2}{(1-u)^2}\sigma_1\frac{du}{d\theta}+\beta_1 i_1=0\tag{17}$$

$$\frac{1}{1-u}\cdot\frac{di_2}{d\theta}+\frac{1}{(1-u)\sigma_1}\cdot\frac{di_1}{d\theta}+\frac{i_2}{(1-u)^2}\cdot\frac{du}{d\theta}+\frac{i_1}{(1-u)^2\sigma_1}\cdot\frac{du}{d\theta}+\beta_2 i_2=0\tag{18}$$

$$\frac{d^2u}{d\theta^2}+2\delta_0\frac{du}{d\theta}+u+u_0=\frac{1}{(1-u)^2}(\Delta\beta_1^2 i_1^2+\Delta'\beta_1\beta_2 i_1 i_2+\Delta''\beta_2^2 i_2^2).\tag{19}$$

3. Equations of Time Evaluation

The equations of time evaluation will be derived from the above equations. In this paper, however, only the equations for $u_0=0$ are presented.

3.1. Operate Motion

Waiting time can be evaluated from Eq. (9). It is obvious that the waiting time is zero when $u_0=0$.

The travelling time can be evaluated from Eqs. (10), (11), and (12). That is, by putting p and i of the following expressions into Eqs. (10), (11), and (12), we can determine the values of A_n and B_n.

$$\left.P=\frac{du}{d\theta}=\sum_{n=0}^{\infty}A_n u^{n/4}\right\}\tag{20}$$

$$\left.i_1=\sum_{n=0}^{\infty}B_nu^{n/4}.\right\}$$

Therefore, the following equation of time evaluation is obtained.

$$\begin{aligned}\theta=&\left(\frac{\beta_1+\beta_2}{\beta_2}\right)\frac{2}{\varepsilon_1}u^{1/4}+\left(\frac{\beta_1+\beta_2}{\beta_2}\right)^2\frac{3}{5\varepsilon_1^2}\\&\times\left(\frac{\beta_1\beta_2}{\beta_1+\beta_2}\right)u^{1/2}+\left(\frac{\beta_1+\beta_2}{\beta_2}\right)^3\frac{2}{15\varepsilon_1^3}\Bigg\{\frac{1}{2}\\&+\frac{49}{40}\left(\frac{\beta_1\beta_2}{\beta_1+\beta_2}\right)^2+\frac{1}{2}\left(\frac{\beta_1}{\beta_1+\beta_2}\right)\Bigg(\frac{\Delta'}{\sigma_1}\\&+\frac{2\Delta''}{\sigma_1^2}\Bigg)-\frac{\Delta''}{\sigma_1^2}\Bigg\}u^{3/4}+\cdots\cdots\end{aligned}\qquad(21)$$

where

$$\varepsilon_1=\sqrt[4]{\frac{4\beta_1^2(\Delta+\Delta'/\sigma_1+\Delta''/\sigma_1^2)}{3}}.\qquad(22)$$

3.2. Short Circuit Release

Waiting time is evaluated by the following procedure. From Eq. (16), the result is:

$$\left.\begin{aligned}i_1&=\frac{\beta_2}{\beta_1\beta_2+\beta_1^2}e^{-\frac{\beta_1\beta_2(1-u_g)}{\beta_1+\beta_2}\theta}\\i_2&=\frac{1}{(\beta_1+\beta_2)\sigma_1}e^{-\frac{\beta_1\beta_2(1-u_g)}{\beta_1+\beta_2}\theta}\end{aligned}\right\}\qquad(23)$$

and from the 3rd equation of Eq. (16), the following expression can be obtained:

$$\beta_1i_1=\sqrt{\frac{u_g(1-u_g)^2}{\Delta+\dfrac{\Delta'}{\sigma_1}+\dfrac{\Delta''}{\sigma_1^2}}}.\qquad(24)$$

Therefore, the waiting time is:

$$\theta_0=\frac{\beta_1+\beta_2}{\beta_1\beta_2(1-u_g)}\log_e\frac{\beta_2}{\beta_1+\beta_2}\sqrt{\frac{\Delta+\dfrac{\Delta'}{\sigma_1}+\dfrac{\Delta''}{\sigma_1^2}}{u_g(1-u_g)^2}}.\qquad(25)$$

Travelling time will be evaluated from Eqs. (17), (18), and (19) by the same procedure Section 3.1.

$$\left.\begin{aligned}P&=\frac{dv}{d\theta}=\sum_{n=0}^{\infty}A_nu^{n/3}\\i_1&=\sum_{n=0}^{\infty}B_nu^{n/3}\\i_2&=\sum_{n=0}^{\infty}C_nu^{n/3}\end{aligned}\right\}\qquad(26)$$

where

$$v=u_g-u.\qquad(27)$$

$$\theta=\frac{3}{A_2}v_1^{/3}-\frac{3A_3}{2A_2^2}v^{2/3}-\left(\frac{A_4}{A_2^2}-\frac{A_3^2}{A_2^3}\right)v+\cdots\cdots\qquad(28)$$

where

$$\left.\begin{aligned}B_0&=\frac{1}{\beta_1}\sqrt{\frac{u_g(1-u_g)^2}{\Delta+\Delta'/\sigma_1+\Delta''/\sigma_1^2}}\\C_0&=\frac{1}{\beta_2\sigma_1}\sqrt{\frac{u_g(1-u_g)^2}{\Delta+\Delta'/\sigma_1+\Delta''/\sigma_1^2}}\end{aligned}\right\}\qquad(29)$$

$$\left.\begin{aligned}B_1&=-\sigma_1\beta_2\sqrt[3]{\frac{3(1-u_g)^4B_0}{\sigma_1(\Delta\sigma_1^2+\Delta'\sigma_1+\Delta'')(\beta_1+\beta_2)^2\beta_2}}\\C_1&=\frac{\beta_1}{\sigma_1\beta_2}B_1\end{aligned}\right\}\qquad(30)$$

$$A_2=\sqrt[3]{\frac{9\beta_1\beta_2u_g(1-u_g)}{\beta_1+\beta_2}}\qquad(31)$$

$$B_2=\frac{\sigma_1\beta_2}{\beta_1}C_2=\frac{\sigma_1\beta_2}{\beta_1}\;\frac{\dfrac{5A_2\beta_1^2(1-u_g)B_1}{\sigma_1C_1(\beta_1+\beta_2)}-\dfrac{\beta_1^2B_1^2}{(1-u_g)^2}\left(\Delta+\dfrac{\Delta'}{\sigma_1}+\dfrac{\Delta''}{\sigma_1^2}\right)}{\dfrac{2\beta_1\beta_2B_0}{(1-u_g)^2}\left(\Delta\sigma_1+\Delta'+\dfrac{\Delta''}{\sigma_1}\right)-\dfrac{10}{3}\dfrac{A_2^2}{C_1}}\qquad(32)$$

$$A_3=-\frac{\beta_1}{\sigma_1 C_1(\beta_1+\beta_2)}\left\{\frac{(\beta_1+\beta_2)}{\beta_1}2A_2\sigma_1 C_2\right.$$

$$\left.+3\beta_1 B_1(1-u_g)\right\}. \qquad (33)$$

3.3. Open Release

Erom Eq. (13),

$$i_2=\frac{(1-u_g)}{\beta_2}\sqrt{\frac{u_g}{\Delta''}}=\frac{1}{\beta_1\sigma_1}e^{-\beta_2(1-u_g)\theta}. \qquad (34)$$

Therefore waiting time is:

$$\theta=\frac{1}{\beta_2(1-u_g)}\log_e\frac{\beta_2}{\beta_1\sigma_1(1-u_g)}\sqrt{\frac{\Delta''}{u_g}}. \qquad (35)$$

Travelling time is:

$$P=\sqrt[3]{9\beta_2 u_g(1-u_g)}\,v^{2/3}-\beta_2(1-u_g)v$$

$$+\frac{9}{20\sqrt[3]{9\beta_2 u_g(1-u_g)}}\left\{\frac{11}{9}\beta_2^2(1\right.$$

$$\left.-u_g)^2-1\right\}v^{4/3}+\cdots\cdots \qquad (36)$$

$$i_2=\sqrt{\frac{u_g(1-u_g)^2}{\Delta''\beta_2^2}}-\sqrt[6]{\frac{9u_g(1-u_g)^{10}}{\Delta''^3\beta_2^2}}v^{1/3}$$

$$+\frac{1}{3}\sqrt[6]{\frac{81(1-u_g)^{14}\beta_2^2}{\Delta'' u_g}}v^{2/3}+\left\{\frac{1}{1-u_g}\right.$$

$$\left.-\frac{1+\beta_2^2(1-u_g)^2}{20u_g}\right\}\sqrt{\frac{u_g(1-u_g)^2}{\Delta''\beta_2^2}}v+\cdots\cdots$$

$$(37)$$

$$\theta=\frac{3v^{1/3}}{\sqrt[3]{9\beta_2 u_g(1-u_g)}}+\frac{3\beta_2(1-u_g)v^{2/3}}{2\sqrt[3]{\{9\beta_2 u_g(1-u_g)\}^2}}$$

$$+\frac{9\{1+\beta_2^2(1-u_g)^2\}v}{20\sqrt[3]{\{9\beta_2 u_g(1-u_g)\}^3}}+\cdots\cdots \qquad (38)$$

where

$$v=u_g-u.$$

4. Force Factor

The value of the force factor of non-polarized relays varies with input current, that is, the force factor of the non-polar apparatus is essentially different from that of the polar one. But in this paper, all the coefficients that have the same form as the force factor of the polar devices will be called the force factor of the non-polar devices in a wide sense.

4.1. General Considerations

The general form of the force factor is Eq. (6); and if eddy current is neglected and the value of u is assumed to be zero, the result is:

$$\Phi_1=\frac{4\pi N_1^2 A}{(AR_0+x_g)^2}I_1. \qquad (39)$$

In the measurement of this factor, mechanical resonance is produced with a frequency double that of the fundamental, and the motional impedance of the fundamental frequency will be given as follows:

$$Z_m=\frac{\omega(\partial L_1/\partial x)^2 I^2}{8s}\cdot\frac{1}{\sqrt{(1-\Omega^2)+(2\omega r/s)^2}}$$

$$\times\frac{\cos(\omega t-\theta)}{\sin\omega t} \qquad (40)$$

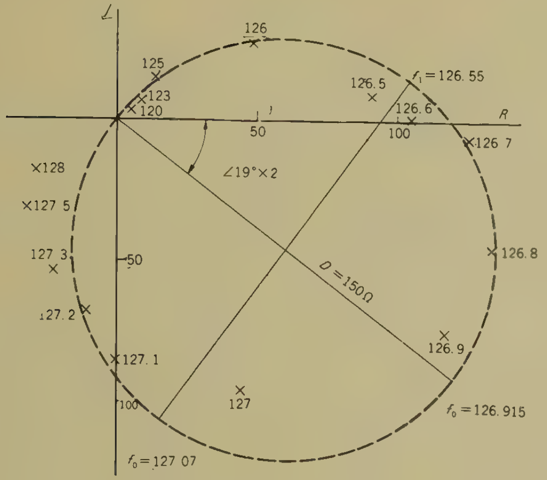
where

$$\Omega=\frac{2\omega}{\omega_0}=\frac{2\omega}{\sqrt{s/m}}.$$

Therefore, the diameter of the motional impedance circle is shown by the sign z_{m0} when $\Omega=1$, and the force factor is:

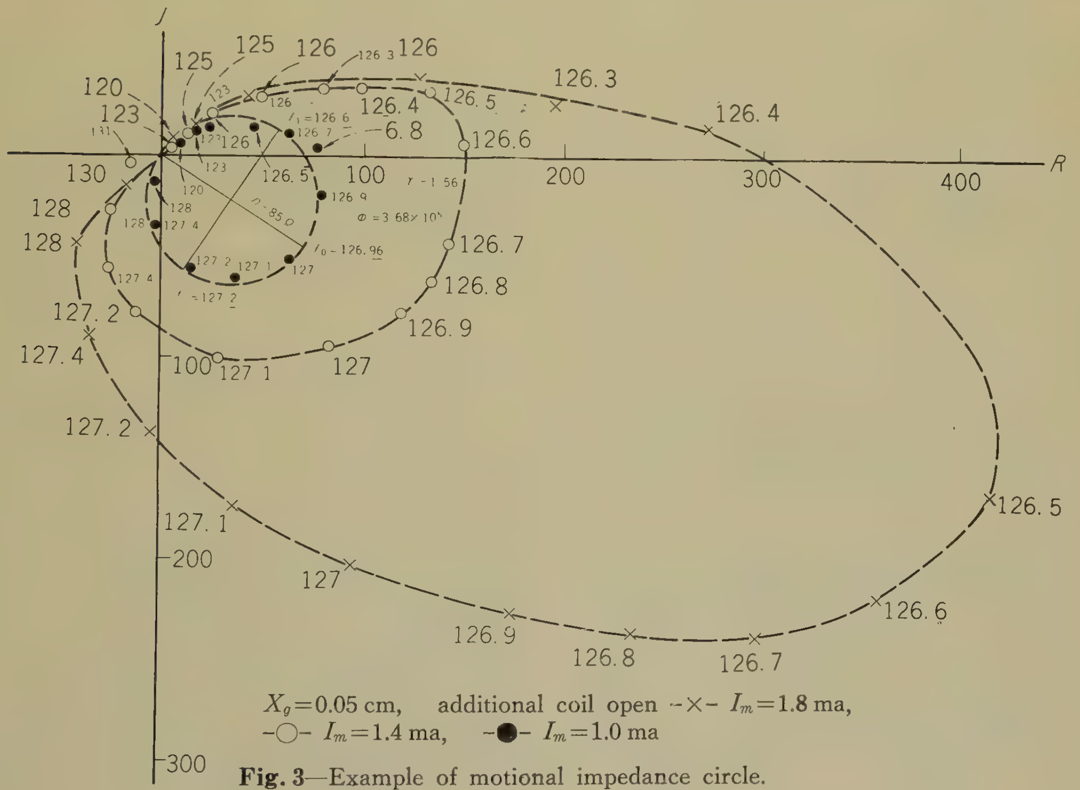
$$\Phi=\sqrt{r\cdot z_{m0}}=\frac{I_1}{4}\left(\frac{\partial L_1}{\partial x}\right). \qquad (41)$$

Therefore the value of force factor shown in this paper is four times larger than that which we measured by small amplitude vibration. As a rule, the diameter of the motional impedance circle is in proportion to the second power of the current, while the



	I_m	r dyne/keine	Φ_1 dyne/ab amp	$D\,\Omega$
a	3	1. 33	4. 46×10^5	150
b	2	1. 15	2. 78×10^5	67
c	1. 5	0. 83	1. 75×10^5	37

Fig. 2—Example of motional impedance circle X_g=0.1 cm additional coil open.



X_g=0.05 cm, additional coil open -×- I_m=1.8 ma,
-○- I_m=1.4 ma, -●- I_m=1.0 ma

Fig. 3—Example of motional impedance circle.

force factor is proportional to the current. Figs. 2 and 3 illustrates this relation.

The relation between the diameter and current can be easily recognized in Fig. 2, and the relation between force factor and current in Fig. 4.

Fig. 3 is an example when the input is large enough in comparison with gap length x_g. The circle becomes elliptic similar to the case of a large output magnet striction transducer.

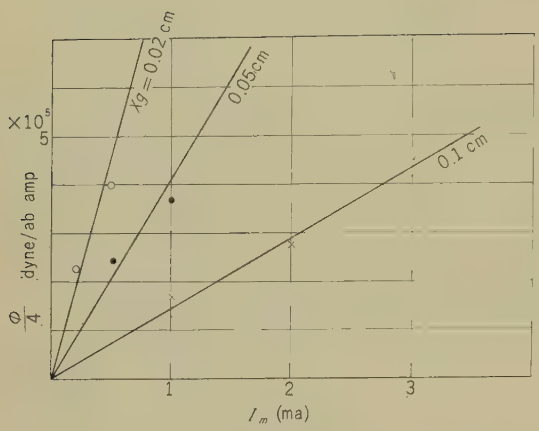


Fig. 4—Relation between the force factor and current.

4.2. Effect of Eddy Current on the Force Factor

In general, the e.m.f. equation of the non-polarized devices is formed similarly to the polar one, that is:

$$\left.\begin{aligned} e_1&=L_{11}\dot{i}_1+L_{12}\dot{i}_2+(L'_{11}i_1+L'_{12}i_2)\dot{x}+R_{11}i_1 \\ O&=L_{22}\dot{i}_2+L_{12}\dot{i}_1+(L'_{22}i_2+L'_{12}i_1)\dot{x}+R_{22}i_2 \end{aligned}\right\} \quad (42)$$

where the prime symbol shows differentiation by x, dot over the letter shows differentiation by time. From the second equation,

$$i_2=-\frac{j\omega L_{12}}{R_{22}+j\omega L_{22}}i_1-\frac{L'_{22}i_2+L'_{12}i_1}{R_{22}+j\omega L_{22}}\dot{x}. \qquad (43)$$

And the force factor is given as follows:

$$\Phi_1=\frac{L'_{11}}{1+j\dfrac{\omega}{\omega_0}\cdot\dfrac{1}{\beta_2}}i_1+\frac{L'_{12}}{1+j\dfrac{\omega}{\omega_0}\cdot\dfrac{1}{\beta_2}}i_2 \quad (44)$$

where

$$L_{12}=\sqrt{L_{11}\cdot L_{22}}\,.$$

As a first approximation, we neglect the effect of $\dot{x}$ in Eq. (43) and put Eq. (43) into Eq. (44). Thus the following equation can be obtained:

$$\Phi_1=\frac{\beta_2^2\left(\dfrac{\omega_0}{\omega}\right)^2-1-2j\beta_2\left(\dfrac{\omega_0}{\omega}\right)}{2+\beta_2^2\left(\dfrac{\omega_0}{\omega}\right)^2+\left(\dfrac{1}{\beta_2^2}\right)\left(\dfrac{\omega}{\omega_0}\right)^2}\cdot L_{11}'i_1. \quad (45)$$

Therefore the absolute value and the phase angle are:

$$\left|\frac{\Phi_1}{L'_{11}i_1}\right|=\frac{1+\beta_2^2\left(\dfrac{\omega_0}{\omega}\right)^2}{2+\beta_2^2\left(\dfrac{\omega_0}{\omega}\right)^2+\left(\dfrac{1}{\beta_2^2}\right)\left(\dfrac{\omega}{\omega_0}\right)^2} \qquad (47)$$

$$\text{phase angle}=-\frac{2\beta_2\left(\dfrac{\omega_0}{\omega}\right)}{\beta_2^2\left(\dfrac{\omega_0}{\omega}\right)^2-1}. \qquad (47)$$

The calculated values are given in Fig. 5. For example, when phase angle is <19° as in Fig. 2, the eddy current constant of this relay is 3. In Fig. 6 are shown other examples with added eddy currents caused by an additional coil.

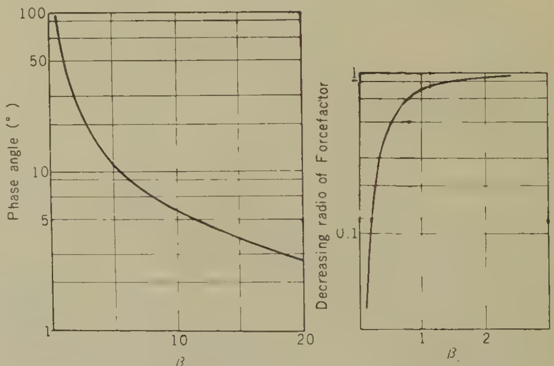


Fig. 5—Effect of eddy current on force factor (calculated value).

5. Method of Evaluation of Constants

5.1. Constants Corresponding to Small Vibration

Constants corresponding to small vibration

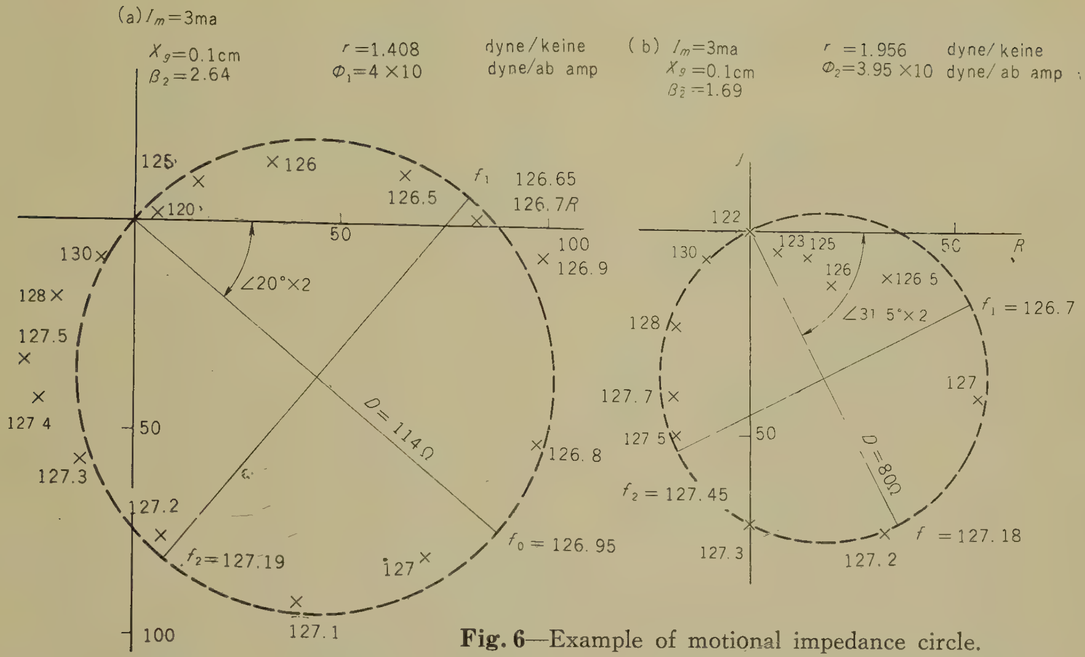


Fig. 6—Example of motional impedance circle.

are impotant for calculation of waiting time, and useful to examine magnetic circuit.

i) When both the eddy current and leakage are neglected,

$$\frac{1}{\Phi_1^2\mu_0}=\frac{R_0}{4\pi sN_1^2}+\frac{x_g}{4\pi sN_1^2A}\tag{48}$$

$$\frac{1}{\Phi_1\mu_0}=\frac{I_1}{s(AR+x_g)}.\tag{49}$$

If N_1 and s are known, we can determine the value of A from the slope of Eq. (48) and R_0 from the intercept of the same equation. But if x_g does not vary, we can determine the constants from Eqs. (48) and (49). But the change of resonant frequency is generally very small; therefore the stability factor is not very exact and therefore not very useful for the non-polarized devices. In this case the following method is preferable.

ii) Eddy current is neglected but leakage is taken into account.

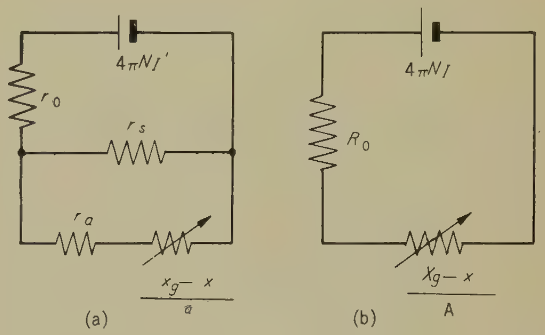


Fig. 7—Equivalent circuit of the structure shown in Fig. 1.

The equivalent circuit of Fig. 1 is given in Fig. 7 (a), and the driving force of the equivalent circuit in Fig. 7 (b) is the same as in (a).

Then:

$$\left.\begin{aligned} &I=I',\quad x_g=X_g,\quad A=\frac{r_s a}{r_s+r_0}\\ &R_0=r_0+r_a+\frac{r_a r_0}{r_s}. \end{aligned}\right\}\tag{50}$$

The effective pole face area A can be determined from the change of the force factor at various gaps.

$$\frac{1}{\sqrt{\Phi_1}}=\frac{AR_0}{\sqrt{4\pi N_1^2 AI_1}}+\frac{x_g}{\sqrt{4\pi N_1^2 AI_1}}. \tag{51}$$

The actual reluctance obtained by actual measurement of the inductance is:

$$R=r_0+r_s-\frac{r_s^2}{r_s+r_a+(X_g/a)}. \tag{52}$$

Therefore

$$\frac{X_g-X_{gc}}{R(X_g)-R(X_{gc})}=\frac{a(r_s+r_a)}{r_s^2}\left(r_s+r_a+\frac{X_{gc}}{a}\right)$$

$$+\frac{X_g}{r_s^2}\left(r_s+r_a+\frac{X_{gc}}{a}\right). \tag{53}$$

The slope y_1 is:

$$y_1=\frac{r_a+r_s}{r_s^2}+\frac{X_{gc}}{ar_s^2} \tag{54}$$

where X_{gc} is the central reference value of x_g. From Eq. (54) and the ratio of the slope and the intercept of Eq. (53), the following equations are obtained.

$$a(r_a+r_s)=\alpha \tag{55}$$

$$\alpha r_s^2=\beta \tag{56}$$

$$y_2=\frac{r_0+r_s}{r_s^2}(r_ar_s+r_sr_0+r_ar_0)+\frac{(r_0+r_s)^2}{ar_s^2}X_{gc} \tag{57}$$

$$\frac{(r_0+r_s)^2}{ar_s^2}=r \tag{58}$$

where y_2 is the slope of the equation

$$\frac{X_{gc}-X_g}{\rho(X_g)-\rho(X_{gc})}$$

where ρ is permeance and $\rho(X_{gc})$ is the central reference value of $\rho(X_g)$. The constants can be determined from Eqs. (50), (51), (55), (56), and (58).

iii) Eddy Current Constant

The value of β_2 can be determined from the phase angle of Fig. 5, and the other constants related eddy current can be written as follows:

$$\left.\begin{aligned}\frac{\Delta'}{\sigma_1}&=\frac{EI_{Is}}{s(AR_0+x_g)^2\omega_0}\cdot\frac{1}{\beta_2}\\ \frac{\Delta''}{\sigma_1^2}&=\frac{E^2}{2s(AR_0+x_g)^2\omega_0^2}\cdot\frac{1}{\beta_2^2(L)_1}.\end{aligned}\right\} \tag{59}$$

5.2. Constant Corresponding to Large Amplitude

When the equation $u_0=0$ is satisfied, the stability factor is as follows:

$$(\mu)_{0\,rel}=\frac{1}{2u_g(1-u_g)^2}=\frac{27}{8}\left(\frac{I_{0pr}}{I_{rel}}\right)^2$$

where

$(\mu)_{0\,rel}$=stability factor at the initial point corresponding to release current.

Therefore the value of u_g can be determined from the measurement of minimum operating and releasing current, and AR_0 can be estimated from the value x_g.

6. Results and Discussion

6.1. Measurement of Constants

The structure of the model was the same as that of the 450 type polar relay, from which, however, the permanent magnet and one of the yokes were removed. At first, the values were estimated to be $m=0.415$ g, $s=10.78\times10^5$ dyne/cm, $f_0=256.5$ c/s by the additional mass method. In Fig. 8 is shown the measured value of inductance, and in Figs. 9–11 values corresponding to Eqs. (53)–(59). The force factor is given in Figs. 2–4. From these figures, the following values can be determined:

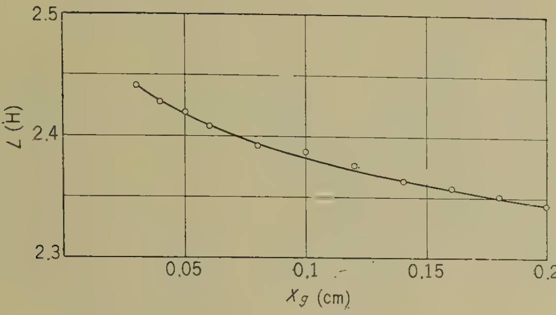


Fig. 8—Measured value of inductance.

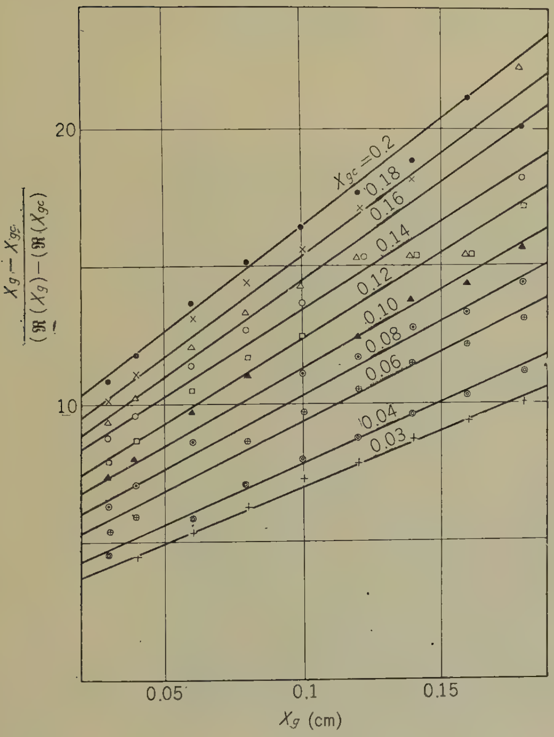


Fig. 9—Relation between gap length and reluctance.

$$a r_s^2 = 0.00477, \quad a(r_s + r_a) = 0.175,$$

$$r_s + r_0 = 0.394, \quad A = \frac{a r_s}{(r_s + r_0)} = 0.1071.$$

Therefore

$$a = 0.373\ \mathrm{cm}^2, \quad r_a = 0.355\ \mathrm{cm}^{-1},$$

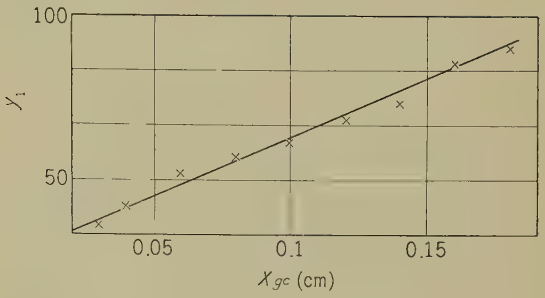


Fig. 10—Relation between the central reference value of X_g and slope of reluctance change.

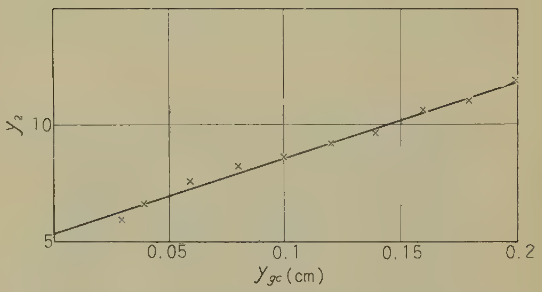


Fig. 11—Relation between the central reference value of X_g and slope of permeance change.

$$r_s = 0.113\ \mathrm{cm}^{-1}, \quad r_0 = 0.28\ \mathrm{cm}^{-1}.$$

Though it may seem that the leakage reluctance is too low, you will see it is a reasonable value if compared with that of the same type polar relay.(1)

The constants which correspond to large amplitude were as follows:

$$I_{0p} = 3.7^{ma}, \quad I_{rel} = 2.75^{ma} \text{ (airgap=0.2 m/m).}$$

Therefore

$$u_g = 0.65, \quad AR_0 = 0.03^{cm},$$

and the value of eddy current constant was 3.

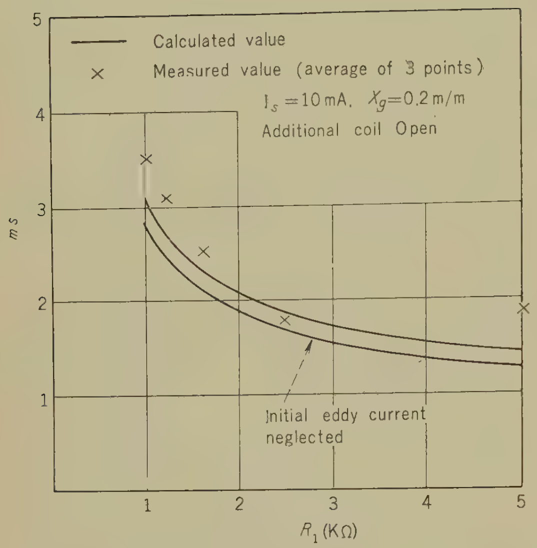


Fig. 12—Operating time₁.

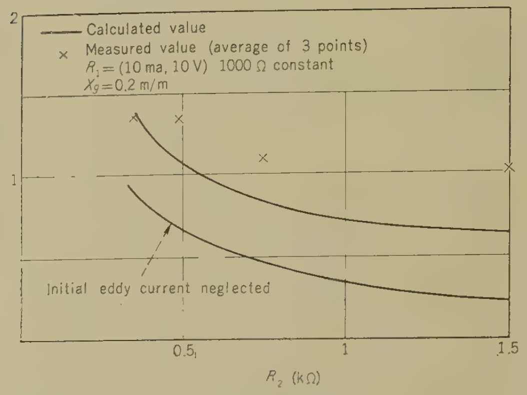


Fig. 15—Open release waiting time₂.

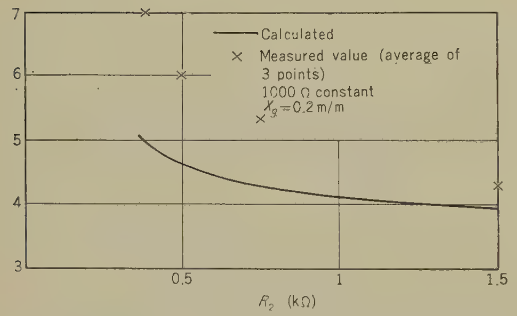


Fig. 13—Operating time₂.

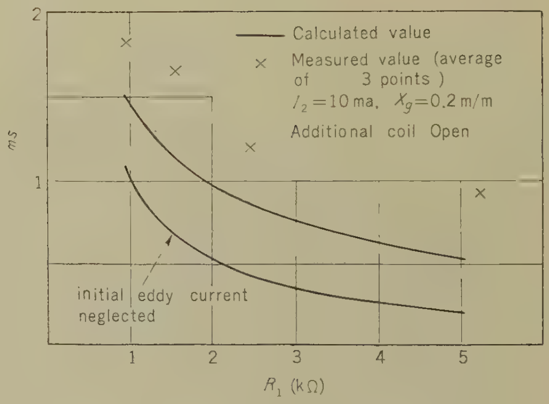


Fig. 16—Short release waiting time₁.

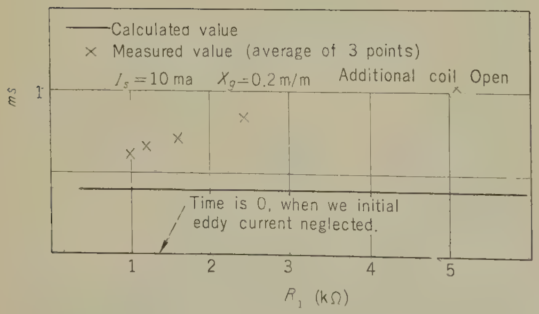


Fig. 14—Open release waiting time₁.

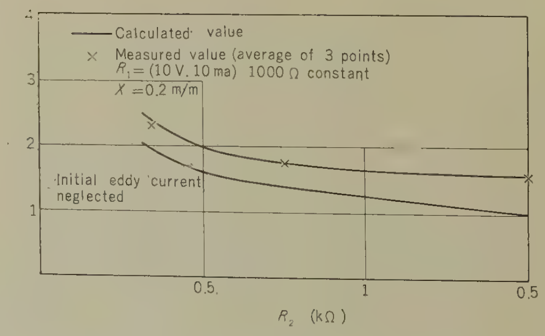


Fig. 17—Short release waiting time₂.

6.2. Measurement of Dynamic Characteristics

Figs. 12 and 13 show the operating time, and Figs. 14 and 15 the waiting time of open release, while Figs. 16 and 17 the waiting time of short release. Fig. 18 shows some examples of the current waveform and the process of its displacement.

6.3. Discussion

So long as it is not so large, eddy current does not exercise much effect on motional time, but this is not the case with waiting time; that is, eddy current has an great effect on the waiting time, especially in the case of open release. These facts can be seen from Figs. 12–17.

7. Conclusion

The motional equations taking eddy current effect of the ordinary relay into account can be derived from the Lagrange-Maxwell equation. The equations of time evaluation was derived from these equations and was compaired with experimental results. A new factor was considered on the measurement for the determination of the coefficient of these equations.

That is, the force factor, to which much attention has not been paid so far in the analysis of the ordinary relay, has been ex-

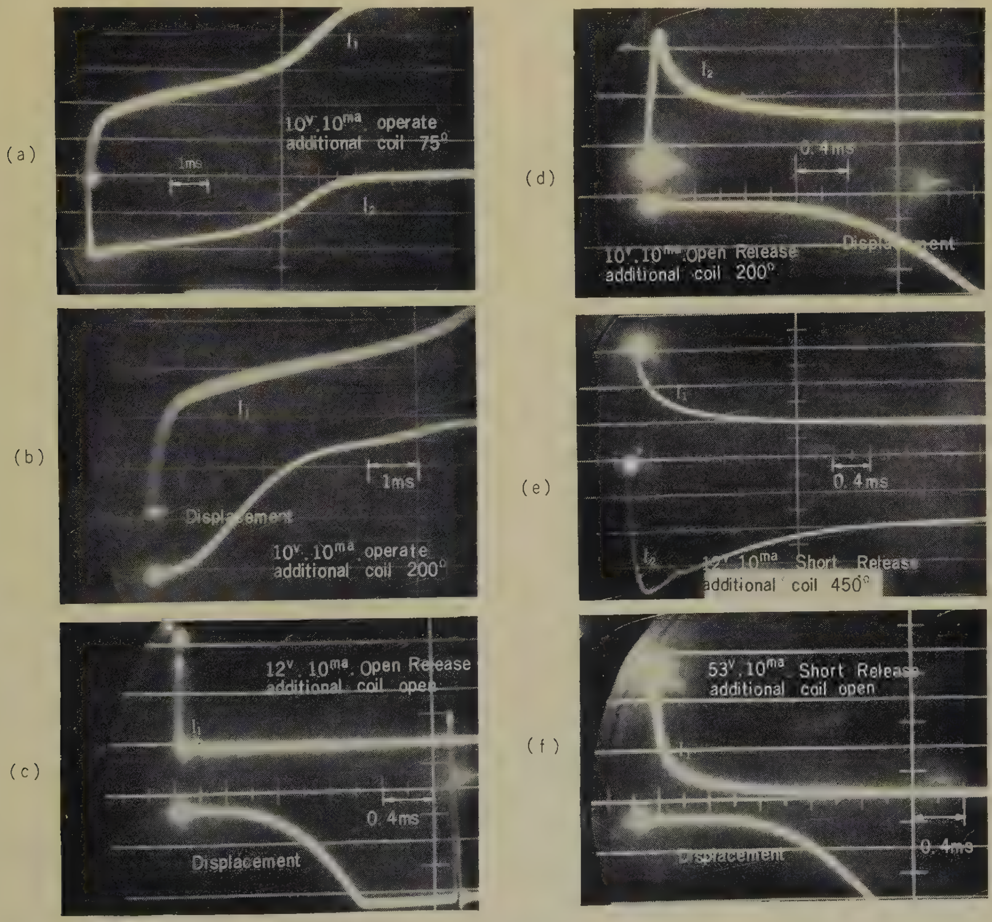


Fig. 18—Current waveform and displacement.

perimentally as well as theoretically shown to be of great importance.

Acknowledgment

The author wishes to express thanks to Dr. Hayasaka and to Dr. Masuzawa for giving him effective guidance.

Reference

(1) Y. Tomita: Analysis of Polar Relay, *Rev of the E.C.L.*, 8, 7/8, p. 297, 1960.

* * * *

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 1-2, JANUARY-FEBRUARY, 1961

U.D.C. 621. 397. 232. 6 : 621. 397. 12

A New Type of Vestigial-Sideband Facsimile System*

Keijiro KUBOTA† and ***Kazuo KOBAYASHI***†

This paper describes a vestigial sideband transmission system for the high-speed transmission of facsimile signals. The system uses simple methods of performing homodyne detection and eliminating quadrature distortion. In the transmitter, the maximum modulation factor is set at a small value, and carrier frequency is constantly transmitted. In the receiver, the even symmetry components around the carrier frequency of the received signal pass through a band-pass filter in which the phase-modulated components are eliminated. Then, the amplitude modulated components are suppressed by a limiter, the output of which is used as the carrier frequency for homodyne detection. Thus homodyne detection can easily be performed without resorting to an A.F.C. system.

1. Introduction

This paper describes a vestigial sideband transmission system designed for the high-speed transmission of facsimile signals. Unlike television, it is not wise to spend much money on equipment for facsimile transmission. Therefore in facsimile, vestigial sideband transmission, if it is used at all, is in most cases received by envelope detection. In such equipment the wave distortion resulting from quadrature distortion(1) is corrected only by a simple circuit.(2), (3)

All these factors have been taken into consideration in the design of the present system, which performs homodyne detection in a simple way, and aims at improving the quality of the received pictures.

2. Principle

Fig. 1 shows the block diagram of the present system and Fig. 2 shows its frequency allocation. In the transmitter, the maximum modulation factor of the transmitter is adjusted to a small value, and carrier frequency is also transmitted during space signals. As illustrated in Fig. 2 (b), the picture frequency band and the carrier frequency are carefully selected so that the whole of one sideband and a part of the other sideband are transmitted through the transmission line. No vestigial sideband filter is used at the transmitter.

At the receiving terminal, the input signal is divided into two parts. One part is fed into the detector after passing through a vestigial sideband filter with the characteristics shown in Fig. 2 (c); while the other part passes through the carrier selection band-pass filter which has even-symmetry response about the carrier frequency. This filter is used to eliminate the phase modulated components of the carrier caused by asymmetrical transmission through the transmission line. The output signal from this filter then passes through a limiter which suppresses the amplitude-modulated components, giving a constant phase and constant amplitude carrier signal,

---

* MS in Japanese received by the Electrical Communication Laboratory, July 13, 1960. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T., Vol. 10, No. 1, pp. 53–70, January, 1961.

† Telegraph Section.

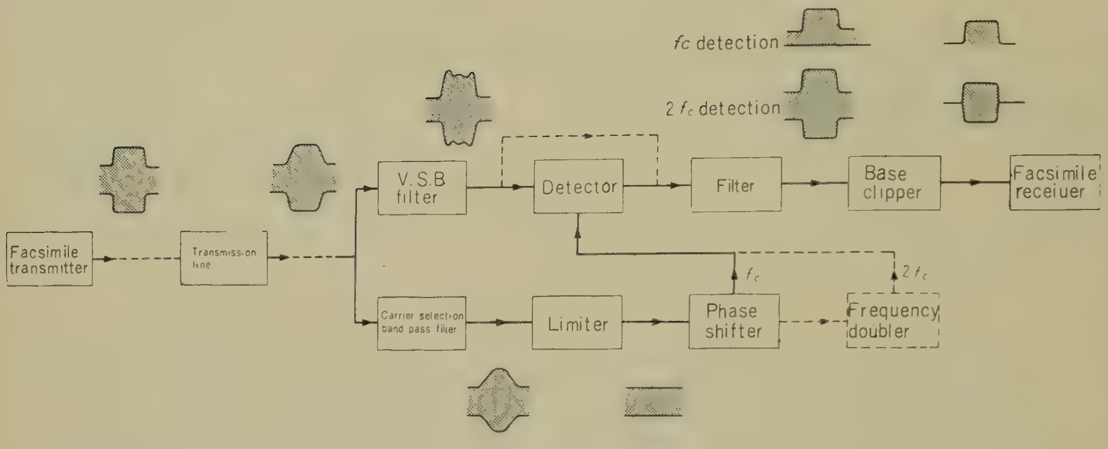


Fig. 1—Block diagram.

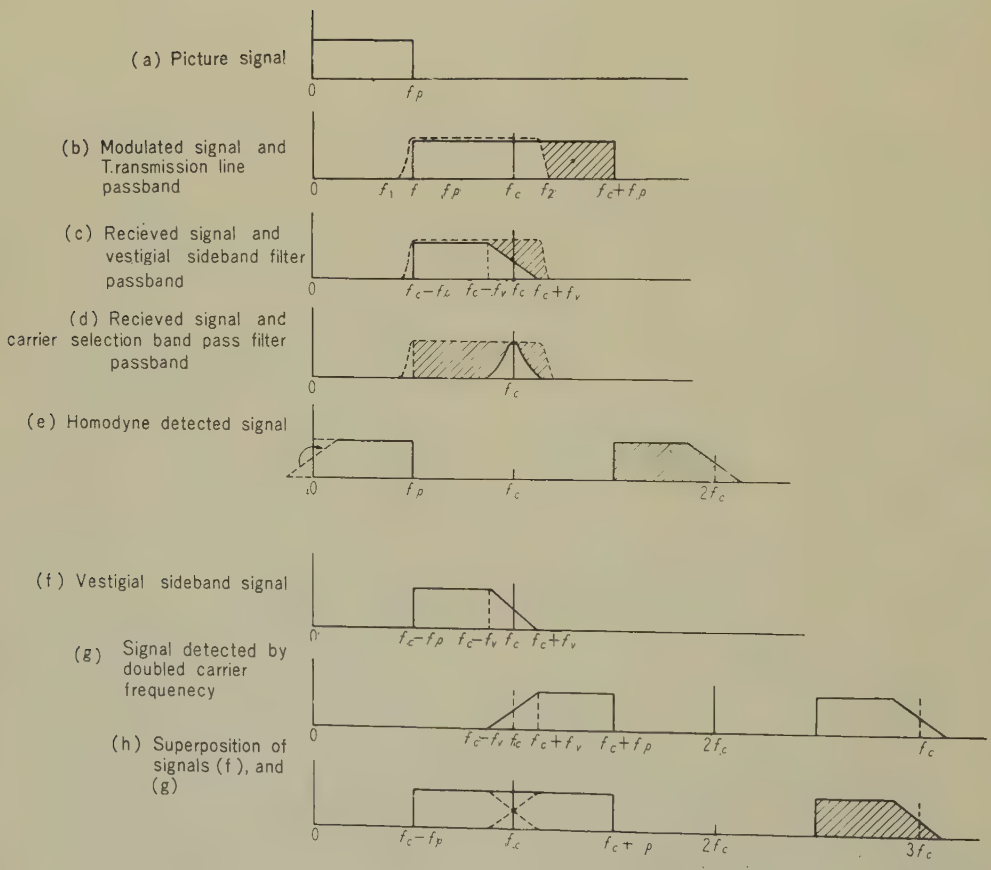

(a) Picture signal
0
fp
(b) Modulated signal and T.ransmission line passband
0
f1
f
fp
fc
f2
fc+fp
(c) Recieved signal and vestigial sideband filter passband
0
fc−fL
fc−fv
fc
fc+fv
(d) Recieved signal and carrier selection band pass filter passband
0
fc
(e) Homodyne detected signal
0
fp
fc
2fc
(f) Vestigial sideband signal
0
fc−fp
fc−fv
fc
fc+fv
(g) Signal detected by doubled carrier frequenecy
0
fc−fv
fc
fc+fv
fc+fp
2fc
fc
(h) Superposition of signals (f), and (g)
0
fc−fp
fc
fc+p
2fc
3fc

Fig. 2—Frequency allocation.

and a phase shifter which adjusts the phase. This constant amplitude, constant phase signal of correct phase is impressed on the detector enabling homodyne detection to be performed. Next, the dc component of the detected signal, which is caused by the low maximum modulation factor, is eliminated by a base clipper; and as shown in Fig. 2 (e), a signal identical to the original picture signal is obtained.

If the equipment shown in dotted lines in Fig. 1 is used for detection and the vestigial sideband signal is superimposed upon the signal detected by the doubled carrier frequency as shown in Fig. 2 (g), a double sideband signal is obtained as illustrated in Fig. 2 (h). It is convenient to use this system when vestigial sideband transmission is performed using double sideband facsimile equipment.

In order to explain the above relation mathematically, the vestigial sideband signal is expressed as

$$P(t)\cos\omega_c t+Q(t)\sin\omega_c t \qquad (1)$$

where $P(t)$=real component
$Q(t)$=quadrature component
ω_c=angular frequency of carrier waves.

If the carrier wave to be used for homodyne detection is

$$2\cos(\omega_c t-\phi) \qquad (2)$$

the output signal of the detector is

$$\{P(t)\cos\omega_c t+Q(t)\sin\omega_c t\}\,2\cos(\omega_c t-\phi)$$
$$=P(t)\,\{\cos(2\omega_c t-\phi)+\cos\phi\}$$
$$+Q(t)\,\{\sin(2\omega_c t-\phi)+\sin\phi\} \qquad (3)$$

If the unnecessary components are eliminated from the above equation, it becomes as follows:

$$P(t)\cos\phi+Q(t)\sin\phi \qquad (4)$$

By reducing the phase angle ϕ to zero, the quadrature component is eliminated and the real component is obtained.

Next the carrier wave to be used for detection by the doubled carrier frequency is expressed as

$$\cos(2\omega_c t-2\phi) \qquad (5)$$

If the same procedure as above is applied, that is, the vestigial sideband signal and the detected signal are superposed and the unnecessary components eliminated. we obtain

$$\frac{P(t)}{2}[\cos\omega_c t+\cos(\omega_c t-2\phi)]$$
$$+\frac{Q(t)}{2}[\sin\omega_c t-\sin(\omega_c t-2\phi)]$$
$$=(P(t)\cos\phi+Q(t)\sin\phi)\cos(\omega_c t-\theta(t)) \quad (6)$$

where

$$\theta(t)=\tan^{-1}\frac{P(t)\sin 2\phi+Q(t)(1-\cos 2\phi)}{P(t)(1+\cos 2\phi)+Q(t)\sin 2\phi}$$

If $\phi=0$, the double sideband signal modulated by the real component can be obtained.

3. Experiments and Results

Since a ready-made facsimile equipment for double sideband was available, the experiments were done according to the system in which detection was performed by doubled carrier frequency.

A signal generator was used as a signal source and the waveforms of the various stages were observed.

Table 1 gives the constants of main parts.

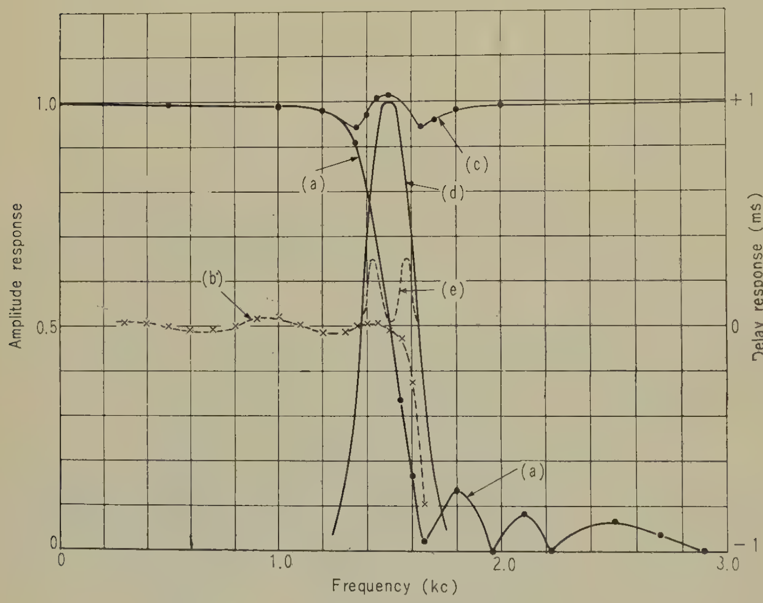
The vestigial sideband filter is a linear phase-shift filter.(4) Its amplitude response is indicated by curve (a) of Fig. 3, and its delay response by curve (b). Curve (c) corresponds to the detected amplitude response.

The carrier-selection band-pass filter selects only the even symmetry components around the carrier frequency among the received signals that have become asymmetrical due to the transmission line. Thus the carrier

Table 1

EQUIPMENT SPECIFICATIONS

Facsimile	Line density : 4 lines/mm Carrier frequency : 1, 500 ∾ Maximum keying frequency : 1, 000 ∾
Artificial Line	300∼1, 800 ∾ band-pass filter
Vestigial Sideband Filter	1, 500 ∾ linear-phase-shift low-pass filter
Carrier Selection Band Pass Filter	1, 350∼1, 650 ∾ band-pass filter



(a) Amplitude response of the vestigial sideband filter.
(b) Delay response of the vestigial sideband filter.
(c) Resultant amplitude response of the detected signal.
(d) Amplitude response of the carrier selection band pass filter.
(e) Delay response of the carrier selection band pass filter.

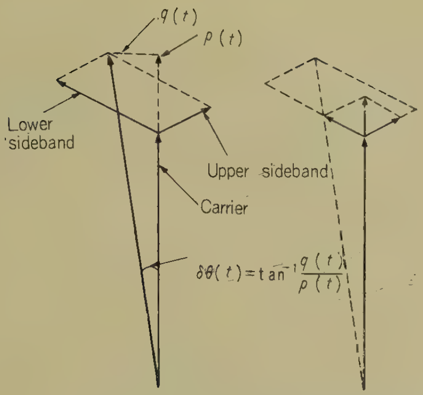
Fig. 3—Characteristics of the vestigial sideband filter and carrier selection band pass filter.

selection band-pass filter serves to eliminate the phase modulated components of the carrier. Its function is explained by the vector diagram in Fig. 4.

The amplitude of the upper sideband and the lower sideband are not equal owing to the asymmetrical sideband transmission. Consequently the carrier wave contains the phase modulated component $\delta\theta$, as illustrated in Fig. 4 (a). By making the upper sideband and the lower sideband symmetrical by the carrier selection band pass filter as shown in Fig. 4 (b), it is possible to make $\delta\theta$ zero. In this case it is necessary for the amplitude response and the delay response of the filter to have even symmetry about the carrier.

As the delay response is not required to be flat, a simple filter can be employed. The amplitude response and the delay response of the filter used in the experiments are indicated by curve (d) and (e) respectively of Fig. 3.

The reason why the vestigial sideband



(a) Received signal (b) Output signal of carrier selection band pass filter

Fig. 4—Vector diagrams of the carrier and sideband.

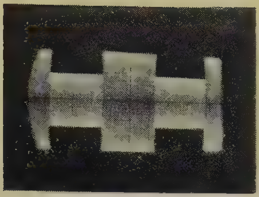
filter is set up in the receiving terminal is to facilitate the selection of the even symmetry components about the carrier frequency.

Fig. 5 shows the waveforms at various points in the present system. Fig. 5 (a) shows the signal generator output, and (b) the artificial line output, in which the distortion because of the asymmetrical sideband transmission is recognizable. Fig. 5 (c) shows the demodulated signal and (d) the waveform after going through the base clipper. It can be clearly seen that the homodyne detection decreases the distortion.

Fig. 6 shows the waveforms of signals with 100 % modulation which were transmitted by vestigial sideband transmission and then detected by homodyne detection using a common carrier oscillator for modulation and detection.

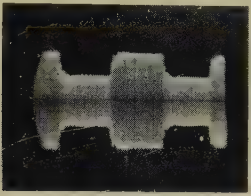
There is almost no difference in the waveforms of Fig. 5 (d) and Fig. 6. This proves the validity of the method of regenerating the carrier at the receiver in the present system.

Fig. 7 (a) is an example of the pictures received by means of homodyne detection with the present system, while (b) is an example of pictures received by means of envelope detection of vestigial sideband transmission. The picture received by homodyne detection



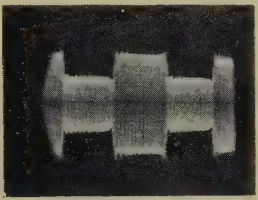
(a)

Signal generator output (Modulation factor: 0. 33, Carsier frequency: 1500 ∾, Pulse width: 10 ms)

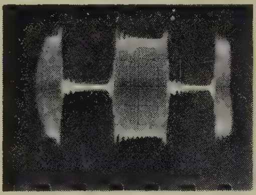


(b)

Received Signal (output of artificial line)



(c) Detector output



(d) Clipper output

Fig. 5—Waveforms of various parts.

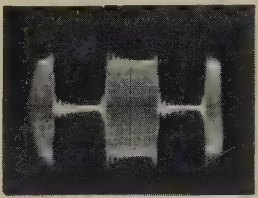
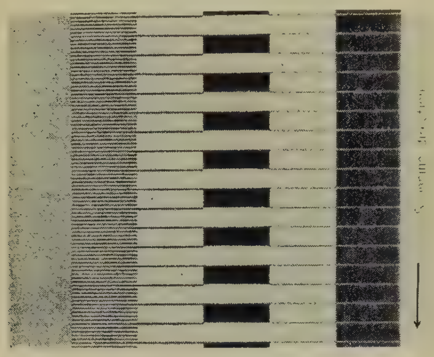
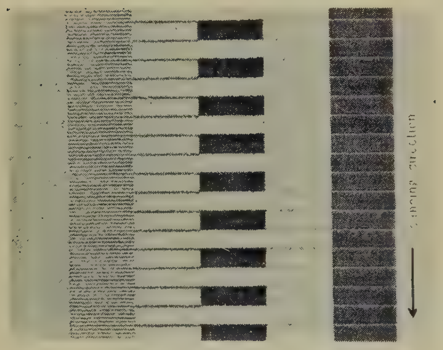


Fig. 6—Waveforms of detector output by carrier oscillator common to transmitter and receiver.

is clearer than that received by envelope detection, and the quality of the received pictures is almost the same as those received by double sideband transmission.



(a) Homodyne detection.



(b) Envelope detection.

Fig. 7—Samples of received pictures.

4. Conclusion

The value of this vestigial-sideband system, in which the modulation factor is set a small value, and the carrier wave constantly transmitted so that it may be utilized for homodyne detection of the signals at the receiver has been demonstrated.

This system is simple and stable, requiring no A.F.C. system for regenerating carrier waves. From the standpoint of the transmission line, the constant transmission of carrier waves is a disadvantage in S/N. But this system is useful for simple and economical vestigial sideband transmission equipment, such as for facsimile communication which utilizes voice lines.

Acknowledgement

The authors of this paper wish to express a heartfelt thankfulness to Dr. T. Kunieda, Dr. H. Okuno, and Dr. M. Kaji who have kindly guided the present study.

References

(1) H. Nyquist and K. W. Pfleger, "Effect of the Quadrature Component in Single Sideband Transmission," *B.S.T.J.*, **19**, 1, p. 63, (1940).

(2) J. V. Hogan, "Facsimile Speedup Enhances Usefulness," *Electronics*, **28**, 7, p. 134, (1955).

(3) E. S. Grimes, "Vestigial Sideband Transmission as Proposed for Intercity Facsimile Signals," *W.U.T.R.*, **13**, 4, p. 152, (1959).

(4) H. W. Bode and B. L. Dietzold, "Ideal Wave Filters," *B.S.T.J.*, **14**, 2, p. 215, (1935).

* * * *

REVIEW OF THE ELECTRICAL COMMUNICATION LABORATORY
VOLUME 9, NUMBERS 1-2, JANUARY-FEBRUARY, 1961

U.D.C. 621. 315. 619. 001. 2 : 678. 652

A Study of Internally Plasticized Melamine Laminates*

Masaaki KATAGIRI†

Paper-base melamine laminates have good electrical properties, good arc resistance, and high hardness; but their punching and postforming characteristics are very bad. In order to obtain melamine laminates which permit punching and postforming processes, the author has studied new melamine laminates made from a high strength cotton paper(1) *with internally plasticized melamine resins. This paper describes the relation between the properties of melamine laminates and the laminating conditions, between the electrical properties and the water absorption, and between the Rockwell hardness and temperature. It has also been confirmed that the internally plasticized melamine laminates obtained by the author have high electric insulation resistance, and excellent punching and postforming characteristics.*

1. Introduction

To improve the postforming characteristics of melamine-formaldehyde laminates, aromatic sulfonamids, amines, and amides have been used, but no good results have been obtained. The first internally plasticized melamine resins were developed by the Monsanto Chemical Company in 1955 (Japanese patent No. 243486, 1958). These resins are melamine-formaldehyde-guanamine condensation products, and laminates made from α–cellulose paper with this resin have good postforming characteristics.

The author investigated various melamine laminates made from a high-strength cotten paper, which he developed, and the following resins: a usual melamine-formaldehyde resin, melamine - formaldehyde - paratoluenesulphonamide (P.T.S.A.) condensation resin (modified by 15–30 % P.T.S.A.), and melamine-formaldehyde-phenilacetoguanamine (P.A.G.) condensation resin (modified by 18.2 % P.A.G.). From these results the author found that the most effective amount of P.T.S.A. for modification was 20–30 %.

2. Manufacture of Laminates

The basematerial and resins used in this study are shown in Table 1. The resin content of the varnish is about 60 % in a solvent which consists of water and methyl alcohol, 4 : 1 by volume. The pick-up is about 60–65 % of dry solids on the weight of the paper. The varnished paper was molded at a pressure of $70\,kg/cm^2$ at 130–160°C, and the molding time was 15–30 minutes. These laminates have a thickness of 1.6 mm.

3. Properties of Laminates

Figs. 1 and 2 show the relation between the properties of the laminates and the molding temperature. At the pressure of $70\,kg/cm^2$ the molding time is 15 minutes. Table 2

---

* MS in Japanese received by the Electrical Communication Laboratory on 30 June, 1959. Originally published in the *Kenkyû Zituyôka Hôkoku* (*Electrical Communication Laboratory Technical Journal*), N.T.T., Vol. 9, No. 12, Dec. pp. 1445–1457, 1960.

† Plastics Applications Research Section.

Table 1

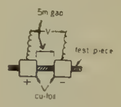
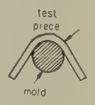
MATERIALS USED

Material	Thickness (mm)	Density (g/cm^3)	Water Absorption (mm/10 minutes)		Tensile Strength (kg/mm^2)		Water-extract Conductivity ($\mu\Omega$/cm) (25° C)
			L.W.	C.W.	L.W.	C.W.	
Base-Material (cotton-paper)	0.099	0.48	48	39	2.2	0.9	29.5

Material	No.	Melamine Resin	Thickness (mm)	Density (g/cm^3)
Resin	(1)	Dimethylol-melamine	1.40	1.47
	(2)	Trimethylol-melamine	1.40	1.48
	(3)	Melamine-Formaldehyde-Phenylactoguanamine (18.2% Modified) Resin	1.37	1.46
	(4)	Melamine Formaldhyde--Paratoluenesulphonamide (15% Modified)	1.44	1.48
	(5)	〃 (29% Modified)	1.45	1.47

Table 2

MEASURED CONDITIONS

Water Absorption (%)	Postforming Properties dia(C_m)	Water Extract Conductivity ($\mu\Omega$/cm)	Electric Corrosion
Test piece ………1.6×25×25 mm	Test piece ……1.6×20×100 mm	Powder………1.5 grs	Method 
Dip in Water ……24 hrs	Heat of test piece ……{175∼180°C 3 minutes	Water-extract ……{water 100 cc boiled 10 min.	
Temp of Water…30±1°C	Mold……………{1∼3 mmϕ metal-cylinder pipe		Test condition
Water Absorption (%) $=\{(W_2-W_1)/W_1\}\times 100$	Bend …………… 	Measured-temp. …25°C	RH−90% 40°C (V)=135 V.DC }96 hrs

shows the measured conditions of the laminate.

The relation between the volume resistivity, tan δ, and the water absorption is given approximately by the following formula:

$$Y=ax^b$$

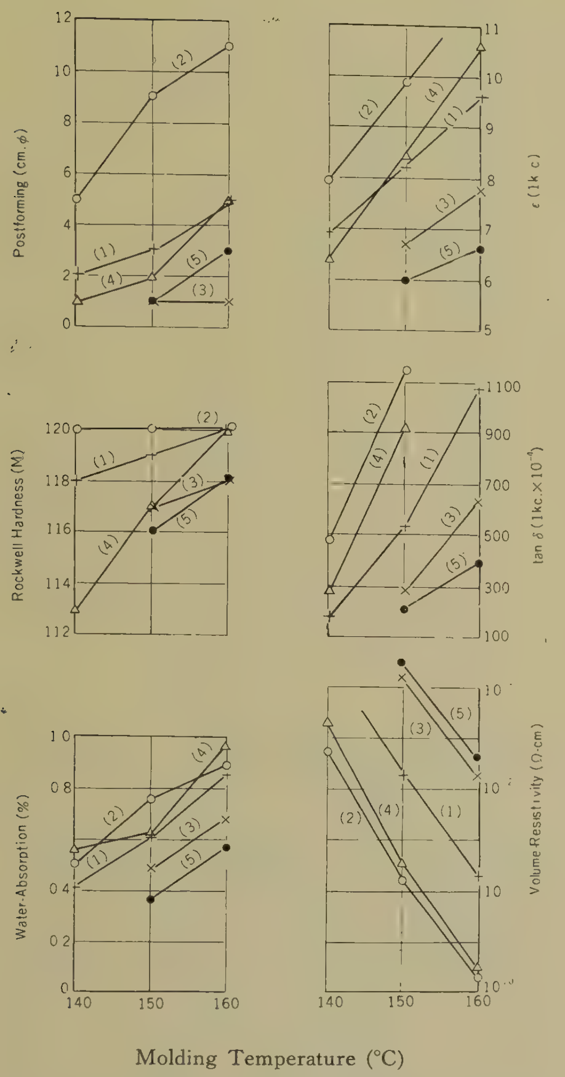


Fig. 1—Relation between properties of laminates and molding temperature.

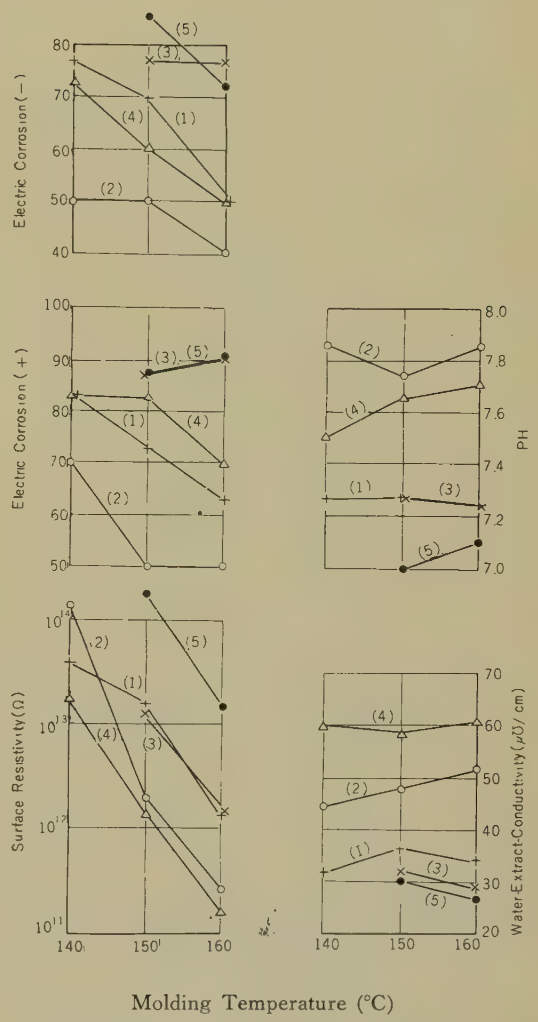


Fig. 2—Relation between properties of laminates and molding temperature.

where Y is the general expression of volume resistivity (ρ_v); tan δ, ε, a, and b are constants; and x is the water absorption (mg/100 cm^2). These relationships are shown in Figs. 3–5.

Figs. 6 and 7 show the relation between the Rockwell hardness and the temperature.

When the molding temperature is 130–140° C, the laminate will burst if subsequently heated as shown in Fig. 6, and the characteristics are very bad; but when the molding temperature is 150°C the properties are very good. However, when the molding temperature is 160°C, the electrical properties become very inferior.

In Fig. 7, No. 5 shows the lowest hardness. No. (3)′ and No. (4)′ are repeated test results of No. (3) and No. (4), respectively.

When the volume resistivity (ρ_v) is plotted vs. the reciprocal of the temperature ($1/T$) an inflection point is observed at the transition point. These results are shown in Fig.

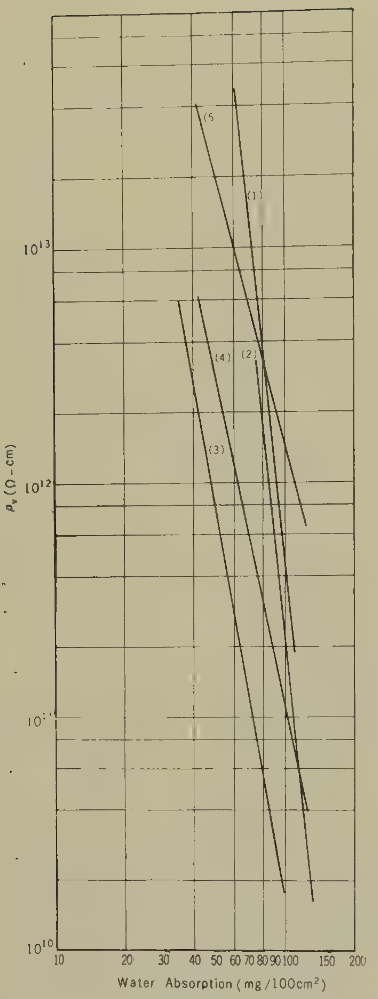


Fig. 3—Relation between ρ_v and water alsorption.

8, where ρ_v is the volume resistivity and T is the absolute temperature. The apparent activation energy of electrical conductivity was calculated from the Arrhenius formula, and these results are shown in Table 3. In Table 3 the H-73 is phenolic laminate.

The general properties of the internally plasticized melamine laminates are shown in Table 4. The laminate which was developed by the author (Sample No. 5) shows and good electric insulation resistance, good punching and postforming characteristics.

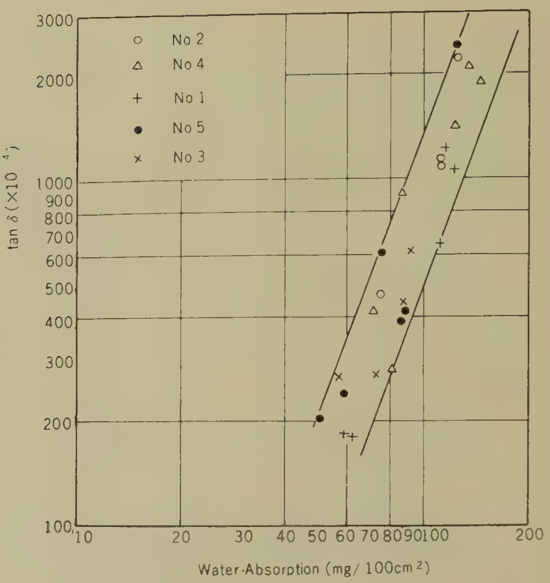


Fig. 4—Relation between tan δ and water absorption.

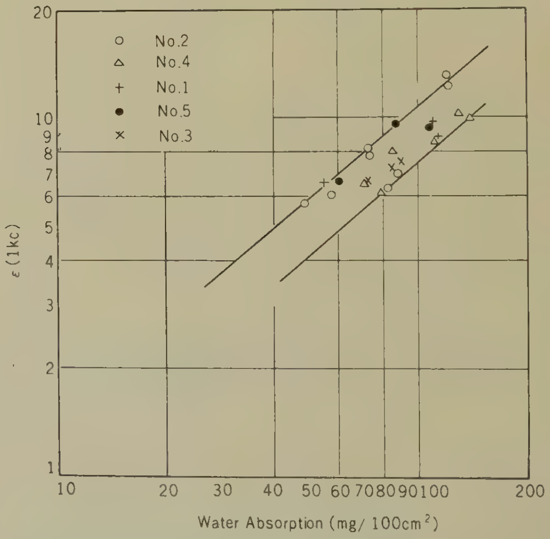


Fig. 5—Relation between ε and water absorption.

4. Conclusion

New internally plasticized melamine laminates made from a high strength cotton paper and melamine-formaldehyde-paratoluenesul-

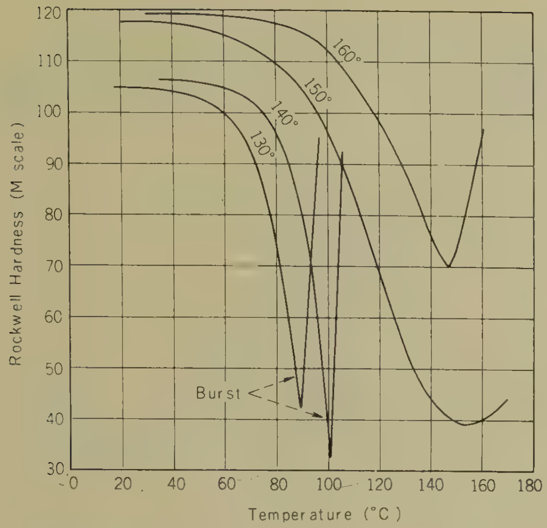


Fig. 6—Hardness vs. temperature (No. 5).

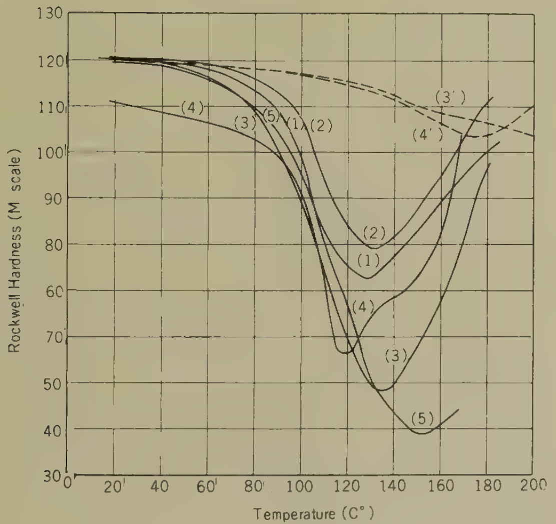


Fig. 7—Hardness vs. temperature of melamine laminates.

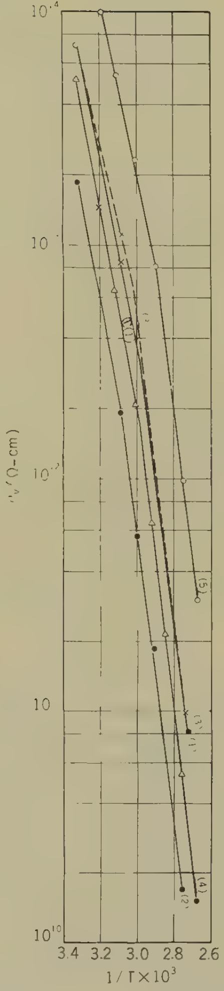


Fig. 8—ρ_v vs. temperature.

phonamide condensation products provide excellent physical and electrical properties when the amount of P.T.S.A. used for modification is 20 to 30%. The less water contained in the laminates the better these properties become. The relation between the electric properties (tan δ, ε) and the water absorption is given approximately by the following formula:

$$Y = ax^b.$$

The postforming characteristics of melamine laminates can be predetermined from the Rockwell hardness vs. temperature curves.

The apparent activation energy of electrical conduction of melamine laminates is larger than that of paper-base phenolics, and the inflection point observed in the ρ_v-temperature

Table 4

PROPERTIES OF MELAMINE LAMINATES

No.	Resin	Base	Thickness (mm)	Rockwell Hordness (M)	Water Absorption (%)	ρ_V ($\Omega\cdot C_m$)	tan δ (1 KC)	Postforming dia. (C_m)
(1)	Dimetylol melamine	Cotton paper	1.42	118	0.41	3×10^{13}	181×10^{-4}	2
(2)	Trimetylol melamine	〃	1.41	120	0.50	3.3×10^{12}	473 〃	5
(3)	M*-F-P.A.G (18.2%)	〃	1.38	116	0.49	1.1×10^{13}	200 〃	1
(4)	M-F-P.T.S.A (15%)	〃	1.45	116	0.56	6.3×10^{12}	285 〃	1
(5)	M-F-P.T.S.A (29.4%)	〃	1.47	116	0.37	2.5×10^{13}	209 〃	1
H-73	Phenol	〃	1.57	98∼105	0.80	4.6×10^{11}	741 〃	–

* M F……Melamine-Formaldehyde.

Table 3

APPARENT ACTIVATION ENERGY CALCULATED FROM FIG. 8

cal./mol.

No.	Temperature		Mean
	Low side	High side	
(1)	23675.0	31998.9	27836.9
(2)	20344.8	32875.0	26609.9
(3)	16814.9	38547.4	27181.1
(4)	19836.6	34284.2	27060.4
(5)	13722.0	29751.8	21236.9
H-73	——	——	19395.8

curve is the second order transition point of resins.

Reference

(1) Laminate Base Materials for Communication Equipment Use Masaaki KATAGIRI; *E.C.L.T.J.*, Vol. 9, No. 5, 125, 1960.

* * * *

Papers Contributed to Scientific and Technical Journals by the Members of the Laboratory

Papers Published in Other Publications of the Electrical Communication Laboratory

U.D.C. 537.226 : 548-162.2 : 537.228.1

Relation between Crystal Symmetry and Piezoelectric Constants of Ferroelectric Polarization.
Kazuo HUSIMI
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **9**, 12, pp. 1421–1434, Dec. 1960

Piezoelectric constants of ferroelectric materials are classified into two parts. A general discussion of these constants is included.

U.D.C. 537.52 : 537.56 : 518.5 : 546.291

Monte Carlo Calculations of the Motion of Electrons in Helium
Tomizo ITOH and *Toshimitsu MUSHA*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **9**, 12, pp. 1459–1472, Dec. 1960

Monte Carlo calculations which were performed with an electronic computer to investigate the behavior of electrons in He gas in a uniform electric field described.

U.D.C. 621.315.619 : 620.193 : 621.318.56.004.6

A Study of the Electrolysis of Laminates used in Communications
Masaaki KATAGIRI
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **9**, 12, pp. 1445–1457, Dec. 1960

The relations between the electrolysis of laminates and the resins of the laminates, the base materials of the laminates, and the electrolytes contained in the laminates are described.

U.D.C. 621.316.541.029.64 : 621.315.212 : 621.317.34

A Method of Measurement of the Reflection Coefficient of Coaxial Radio Frequency Connectors
Takashi INATOMI
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **9**, 12, pp. 1435–1443, Dec. 1960

A study on the methods ot measurement of the reflection coefficient of connectors, including the results of a method developed by the author is reported.

Translations of many of these papers will appear in a future addition of the Review of Electrical Communication Laboratory.

Please address requests for copies of these papers, which are in Japanese, directly to the authors to whom credited at: The Electrical Communication Laboratory, 1551 Kitizyôzi, Musasino-1 Tokyo.

U.D.C. 621.318.56.066.6.001.2 : 621-272 : 531.66

Chatter Vibration of the Switching Relays
Masao TAKAMURA, Toyoshi YURINO, and *Yuji OTUKA*
Kenkyû Zituyôka Hôkoku (Electr. Comm. Labor. Techn. Journ.), NTT, **9**, 12, pp. 1355–1420, Dec. 1960

An analysis of the physical mechanisms which control the chatter of switching relays used in communications apparatus or equipment is described.

* * * *

Vol. 9, Nos. 1-2 JANUARY-FEBRUARY

CONTENTS

On Fundamental Qualities of Vowel Sounds in Speech Communication (Ⅲ) 1
Toshiyuki OCHIAI and *Teruo FUKUMURA*

Staggered Operation of Doubly Resonant Parametric Amplifiers 20
Shigetaka HAMADA and *Hisakazu MUKAI*

Construction and Characteristics of Silver-Bonded Diodes 26
Shoichi KITA

Determination of Molecular Distribution of High Polymers by Fractionation 31
Hiroshi OKAMOTO

Core Matrix Driver Employing Magnetic Amplifiers 43
Junji YAMATO and *Yasunobu SUZUKI*

A Consideration on Relation between Mechanical Behavior of Glasses and Their Internal Structure 50
Muneo WATANABE and *Taro MORIYA*

Effects of the Force Factor and the Eddy Current on the Operate and Release time of the Ordinary Relays 72
Yasuo TOMITA

A New Type of Vestigial-Sideband Facsimile System 85
Keijiro KUBOTA and *Kazuo KOBAYASHI*

A Study of Internally Plasticized Melamine Laminates 91
Masaaki KATAGIRI

Papers Contributed to Scientific and Technical Journals by the Members of the Laboratory 97

Copyright 1961 by the Electrical Communication Laboratory, NTT